



The Alberta Journal of Educational Research

Vol. XII, No. 2

June, 1966

CONTENTS

	PAGE
Student Values as Educational Objectives <i>M. Scriven</i>	89
Teachers and Educational Decision Making <i>A. Kowalski and W. D. Knill</i>	105
Methods of Predicting Freshman Success: Summary and Evaluation <i>D. B. Black</i>	111
Differentiating Prospective Elementary and Secondary School Teachers <i>O. H. Bown, F. F. Fuller and H. G. Richek</i>	127
Methods of Teaching Reading to Beginners <i>M. H. Neville</i>	131
Knowledge of Science and Teaching Behaviour <i>P. J. Eccles</i>	141

COMMITTEE ON EDUCATIONAL RESEARCH
The University of Alberta

FORTCOMING ISSUES

Vol. XII. No. 3, September, 1966

SPECIAL ISSUE

MODELS IN RESEARCH AND THEORY.....

Articles by:

Elizabeth and George Maccia, The Ohio State
University

Marc Belth, Queens College, City University of
New York

Charles J. Brauner, University of British Columbia

Ronald Jones, The University of Alberta
and others



AN INVITATION TO CONTRIBUTORS

A special issue devoted to replications, comments and criticisms of the work of Jean Piaget and his associates is being planned for publication during 1967. Scholars interested in this field are invited to submit manuscripts or give notice of intention to submit. The date of publication will be determined by the availability of sufficient suitable copy.

Editorial Comment

THEORY, RESEARCH AND PRACTICE

A great deal has been written about the gap between educational theory and educational practice. Sometimes reference is to the time lag between the propounding and development of a new point of view with respect to the nature of the educative process and the attempts to work out and apply that idea in educational institutions. Sometimes the reference is to the discrepancy between the latest findings of basic research in learning theory and their applications in classrooms and lecture halls. The former interpretation, although closely related to the latter, is less relevant for the pages of a research journal than the latter on which attention is concentrated here.

It is not proposed to establish that there is a significant "gap" but to assume it and point to some of the reasons and some of the ways in which it might be narrowed. Analogies are weak and often dangerous bases on which to rest an argument and generally, refraining from using them, except to initiate understanding—not explanation—is wise, but there is one that illuminates the difference between the research-practice relationship in education and the same relationship in another applied science field.

Let us envisage the stages through which an aeronautical engineer might be assumed to pass during the development of a new airliner and contrast this with the process which is more typical of educational research. The engineer starts from a set of specifications or requirements with respect to speed, pay-load, operational efficiency, range, cost and so on, that are provided by the potential users. Using these as his criteria he begins to operate mathematically and theoretically to develop one or more possible designs and to calculate the aerodynamic factors that determine the structure and form of the materials and components necessary. The process is wholly theoretical at this stage and the end product is a drawing board version of the craft. The development and testing of these components is, initially, a laboratory process but a stage is reached when practical tests must be undertaken to check and correct the laboratory calculations. Metal strength and its fatigue characteristics can be estimated in the ideal laboratory situation but the final test must be carried out under practical conditions. At the same time

miniature isomorphic models* are constructed and used to test aerodynamic factors in wind-tunnels and other laboratory devices. Such tests are used to modify the initial design and they do contribute significantly to modification of it but they go only part of the way. Full scale mock-up models are built to test structural details and appointments. Prototypes are built and exhaustively tested under conditions parallel to those likely to be encountered by the production version. Weakness in the *theory* that has guided the design are discovered and corrected. Modifications are introduced as a result of tests in these idealized practical situations. The production model is carefully tested prior to delivery—prior to full scale production.—and modifications are made. As soon as the production model goes into service an intensive follow-up study is undertaken and, if necessary, more modifications incorporated. In all these stages practice is being used to correct and refine the theory and the research on which it is based. There is no need to labour the point. Readers can draw the conclusions relevant to the relationship between research and practice in education. The relationship between research and practice (or between theory and practice) is cyclic. Practice suggests areas of needed research; research suggests theory which in turn delineates research; research recommends to practice and practice corrects, refines and modifies the findings of research and the theories formulated.

If the research-theory-practice relationship is conceived in this way the tendency to conceptualize the theory-practice relationship as a “theory —→ practice” dichotomy with practice in a subservient and passive role is likely to disappear. The result is the conception of practice as an integral aspect of research with a positive guiding role. It necessarily follows that educational institutions and personnel have to be involved in the research enterprises. How can this be achieved? Several things seem to be so obvious that mention of them is scarcely necessary.

First, there is in our schools and colleges a large untapped reservoir of interested and talented teachers waiting to be given the chance to participate effectively in the search for better methods and forms of educating. They need time to read, time to think and time to experiment. The time will not be available to them as long as the administrative system demands the same *kind* of service to the *same* degree from all. This applies to all teachers—elementary, secondary, technical and university. Teaching is important, research is important and publication is important as are the other kinds of service that the professional teacher can be

*Isomorphic miniatures are very simple models with limited explanatory value. However, their use in the educational contest is long overdue. There are few, if any, schools or universities that do not possess serious impediments to ingress and egress of students at peak periods and yet architects seem not to have considered the possibility of determining these limitations and modifying their designs by constructing models and carrying out necessary research before the building is erected.

called upon to give. They are not all equally important for all and the decision as to which, for any particular individual, is the most important can be made only on the basis of the interests, capacities, technical skills and propensities of that individual. An enlightened system requires that each contribute according to his ability. If employers and administrators will accept this principle and its implications then teachers will also accept that one will stress one aspect and another some different aspect.

Second, teachers in schools, colleges and universities cannot be expected to contribute significantly to research, or any other area, unless they are given adequate assistance with the technical problems of design and analysis. Universities and research institutes have personnel competent to provide this assistance. They need to be both encouraged and expected to take a positive position. This is not a responsibility that can be discharged by the staffs of the faculties and colleges of education alone nor should these persons be expected or allowed alone to do so. All university professors, and particularly those in the basic areas such as the arts and sciences, need to become truly concerned with the teaching of the disciplines which they profess and involved in substantial research into that teaching at all levels. Too many professors in arts, science and education are too willing to destructively criticise the effectiveness of the elementary and high school teaching methods and procedures and too unwilling to inform themselves about that teaching. Participation by a significant proportion of these people in research programmes designed to reveal the relationship between what is currently being done and what in fact can be done is likely to narrow the discrepancy between the two and to increase the awareness of all of the problems facing others. Nothing but improvement can result from involvement of university teachers of a discipline in research programmes in co-operation with high school teachers of the same discipline.

Thirdly, university summer sessions are a distinctive characteristic of North American universities. The students enrolled in these courses are predominantly teachers. This appears to be a healthy situation. Unfortunately very few of the enrolled students are concerned with anything more significant than the completion of an undergraduate degree. They are enrolled because they need and are expected to complete their professional education. That this is recognised by the school systems which employ them is shown by salary increases consequent upon completion of a university degree. There are, at the same time, school systems which prefer to employ partially qualified teachers because it is financially expedient to do so. The argument that opposition to delaying teacher certification until university graduation stems partly from the desire to maintain a supply of "cheap" teachers can be empirically supported. Also, a case can be made for suggesting that this opposition would lessen if the lay members of school boards could not argue (rationalize) that completion of full professional qualifications prior to accepting an appointment is not essential because summer and evening

courses facilitate "on the job" completion. University administrators and academics need to consider carefully and seriously whether the current emphasis on the provision of these courses in quantities adequate to fill the present demand is a service or a disservice to teachers and to education. I would be the last to suggest that any impediment be placed in the way of any person who wishes to benefit from further education. I would also be the first to suggest that a system which *requires* a large proportion of its professional personnel to discharge their professional responsibilities by day and complete their *basic* professional qualifications by night or by summer is an unjust system. It is more than unjust—it is morally wrong. It is not expected of professions less important than teaching.

This does not mean that immediate termination of courses of these types is recommended but it does mean that more consideration be given to progressive reduction, after due notice, to encourage more students to complete their basic professional education before entering the profession. Reduction of the summer burdens which now fall heavily on some university faculties and many academics will free them in the evenings and during the summers to offer specialized courses for qualified teachers who wish to proceed beyond the minimum. We need more highly qualified teachers for teaching and for research. Undergraduate programmes cannot provide for this need even if all teachers complete a degree before beginning to teach. Summer time can be more advantageously used to provide specialized institutes for the improvement and up-dating of subject-matter knowledge, the development of research competencies and familiarization with new discoveries and techniques. The need for special training in research methods is urgent but even more urgent is the need to involve a significant proportion of teachers in research projects some of which can be continued as part of the overall programme during the academic year following the summer institute. This will lead to the fully integrated relationship between research and practice that is essential if the latter is to perform its positive role in the discovery of improved educational methods and forms.

Desirable as summer research institutes may be they must be viewed only as one possible minor aspect of the solution. For many, more extensive courses are necessary and the possibility of financial support to provide release time for teachers who have the ability to profit from specialized graduate diploma programmes needs to be investigated. However, until this can be done, if each summer only a mere handful of teachers, selected by their ability and interest, can be involved in a significant research programmes under competent direction and given some understanding of the relevant techniques they are likely to go back to their professional positions able and willing to make some applications and hence provide some data for the correction of theories evolved in the "idealized" laboratory situation. This may lead toward creation of the desirable situation in which practice refines and facilitates research and theory.

The author discusses philosophical and methodological problems associated with the determination of appropriate values and their validation, and considers the practical implications for teaching and testing values. The emphasis given to methodological questions related to the empirical determination of values and the measurement of values should prove interesting and valuable for a wide range of researchers. The author argues for the social necessity and the logical defensibility of the inclusion of teaching about values (not indoctrination into values) in the curriculum and concludes with some comments about the extremely poor level of research on value measuring instruments.

MICHAEL SCRIVEN

Indiana University

Student Values as Educational Objectives*

Introduction

Treatment of this topic raises four problems, two in the philosophy of education and two in its methodology:

1. Can one *justify* trying to change student values at all? In particular, would this make education propaganda, or would it violate some inviolable facts/value distinction?
2. Can one justify *one particular* set of values towards which one should direct one's students? In particular, what nontrivial values

*Professor Scriven will welcome comments and criticisms of this paper from educators and psychologists and particularly from anyone interested in research into values, value theory and the teaching of values. He has specifically asked that this invitational note be appended. (Ed.)

Reprinted from the *Proceedings* of the 1965 Invitational Conference on Testing Problems with permission of the publisher, Educational Testing Service, Princeton, New Jersey.

could legitimately be advocated in the secular state schools of a pluralistic democracy?

3. Can one ever demonstrate the occurrence of changes in student values to the educational factor? In particular, how can one handle the control group problem (a) within the constraints of social taboos on withholding education, (b) in face of the relevant analogue of the Hawthorne effect, (c) given the long time-scale probably involved in a significant value change with the consequent interference from many independent maturational variables.

4. Even more fundamental, can one really *measure* student values in any important sense? In particular, can one get past the superficiality of catalogues of announced values, and past the dubious inferences from projective tests, to a reliable construct of the response-tendencies that would dominate behaviour in a real, conflictual, value-loaded situation?

Each of these questions is of very great importance and difficulty, as is attested by the vast magnitude of the relevant literature and the marked triviality of much of it. Fortunately, the limitations of length on this paper absolutely guarantee *its* superficiality and thereby mask its author's shortcomings. In view of the topic allotted to my distinguished colleague on the panel at this session, I shall concentrate on the two philosophical questions, although I shall say a word or two about the others. All I can really do is to state my view of the present research situation in these fields, with a brevity that will surely seem dogmatic, but with a configuration of reasons that you may find interesting. In the background of my thinking, of course, is the vast sea of the values literature, particularly the strong current containing Philip Jacob's book, *Changing Values in College*, and the discussion of it, especially the one by Barton together with his definitive elaboration of the general issues in his paper, "Measuring the Values of Individuals"; and of course the existing armoury of tests, which we might describe as the *Fifth Mental Measurements Yearbook* plus appropriate (imaginary) supplements. On the philosophical side, I am thinking not only of what might be called the 100 Years War in ethics, but of the long history of debates about the role of moral education in the American school system and most recently the attempts by Hunt and Metcalf, Don Oliver, the World Law Fund, the Ethical Culture schools and others to justify or implement such practices. One could hardly do justice to even one of these projects in a few minutes of comment, and so I shall do an injustice to the subject as a whole instead.

My answers to these four questions require a one-sentence preamble. (However, it may be the most complex one-sentence preamble you have ever encountered.) We must distinguish values in the very widest sense, which includes standards of any kind referring to any field (preferential values), from moral (normative) values and these

in turn from personal standards of behaviour and thought (prudential and conventional values); and we must distinguish between the widest *spread* of the term "value", which includes every item-preference, and the sense in which it refers to more abstract criteria (honesty, etc.); and we must distinguish objective values (if any) from (a) falsely professed values, (b) truly professed values, (c) truly professed and actually operative values, and (d) implicit values*; and we must distinguish values in the sense of external goals from values in the sense of internal sets or attitudes, and values as individual properties from values as group properties. The reason we must make these distinctions is not that they have any effect on the answers, because they do not, but simply that if we don't, someone will be sure that by overlooking them we have invalidated our arguments. The answers themselves also require only one sentence, and if *only* we could say as much (or as little) for their justification, we could all go home early. Of course, there can't be anything very novel about anyone's answers to these questions since all possible sets of answers have been given many times. The novelty, if any, must lie in the particular procedure of justification. Indeed, what is most striking about these issues is the number of bad arguments that are adduced for correct conclusions. To answer the questions then: we *can* frequently justify trying to change student values in the classroom; we *can* frequently justify trying to change them in a particular direction; we *can* in fact measure them and change them and show that we do.

These answers have a nice nineteenth century naiveté about them, and the sharp specialist is likely to view them with suspicion although probably a majority of us, at least *secretly*, believe that they are the right answers. What has led specialists to suspect these answers? Mainly the profound difficulties that have emerged in the field of ethics with the utilitarian and naturalistic approaches, in the field of tests and measurements with validation of the allegedly value-sensitive instruments, and in the field of experimental design with the analysis of long-term, *ex post facto* designs. I propose to say a very few words about each kind of difficulty but considerably more about the philosophical issues than the methodological ones.

The "Value-Free" and "Value-Involved" Positions

A great deal of discussion has gone on between the advocates of "value-free" social science and what we might call the "value-involved" school. The discussion has been extraordinarily deficient in the kind of painstaking examination of simple examples that correspond in logic to the collection and study of extensive data in the empirical sciences. Perhaps this is because the Wittgensteinian revolution in philosophy mainly

*Values that their owners reject but which nevertheless motivate them.

consists in the realization that there is an activity of "logical data-collecting" which is absolutely fundamental to the formulation of logical theories. *A priori* theories are just as inappropriate in logic as in science, even though a logical theory itself will, and a scientific one will not, be *a priori* in another sense.

If we do examine the kind of value judgements that are often fully supportable in the consumer research field, for example, and the kind of recommendations that are often justified to the hilt in the medical field, we begin to smell a rat in the value-free story. It seems clear enough that certain value judgments can be backed up all the way back to breakfast, *using* the resources of various sciences and technologies. Whether we call the propounder of such claims a scientist, an engineer, an executive, a consultant, a philosopher, or a technician seems singularly insignificant. If it is appropriate for a public health officer to condemn certain food-processing practices on the basis of extensive bacterial and other tests, and in terms of a set of criteria based on many years of field experience by many other doctors, it seems less interesting to argue about classifying his job than to recognize his action as the production of a scientifically warranted value judgment. And his condemnation is not intrinsically different from condemnations by specialist anthropologists of early practices by the Bureau of Indian Affairs, by psychologists of certain uses of the IQ tests, by penologists of the death penalty, by political scientists of the form of government of the District of Columbia, by labour sociologists of pure piece-work wage structures, by economists of tariff raises as a protective device for a domestic industry, or for that matter, by physicists of certain interpretations of quantum theory. Criticism and approval are a necessary part of the process of internal improvement of a science, as well as of the process of external application of science, and value judgments expressing them are important and complex, and hence much debated, but absolutely inescapable—except by the ostrich route.

Of course, value judgments do not spring full-fledged from *the facts about the entity being evaluated*, but that does not show they are not empirical. They require a careful combination of those facts with other *facts about the needs, wants, and ideals of the valuing agents*.

It is perfectly true that the codification of public health standards is greatly simplified by the remarkable *universality* of the desire to avoid ill health, which is quite distinct from its *validity*. (I will discuss validity in a moment.) The basic value which the public health officer combines with his bacterial counts and sampling procedures, i.e.—with his specifically scientific skills, is rather more *easily* identified as the chief relevant value than those which bear on the great disputes in the social fields. All that follows from this is that the value judgments with which the social scientist is concerned need more of his professional skills before

they can be substantiated. For it is *his* skills that are required to identify the wants and needs and ideals of the people concerned. The value-free protagonist imagines that the scientist's task ends with presenting the facts about the alternatives being evaluated—let us say, about the use of advanced placement tests or new biology curricula. Then, on this account, the decision-maker selects in accordance with his own values. What an abrogation of professional responsibility this is! It is simply a fancy dress version of the old buck-passing and fee-splitting games. Suppose that the market research and package designing firms hired by a cereal manufacturer to face-lift its line were to come up with seven different packages and the comment that these appeal to various groups, and the manufacturer can just make his own decision. But how does he make it, if their groups do not happen to coincide with the ones of interest to him? They would have done either half or twice the job for which they were hired. It is up to *them* to find out, at the *beginning* of their research, which market is being attacked, or what parameters are to be used if a combination of submarkets is to be invaded. And they won't get this answer in that form by just asking for it; they'll probably have to ask a great many questions, construct hypothetical situations, look at the power structure of the firm in order to be sure that they're getting the answers from the source that will make the decision, etc. And they'll probably have to re-do some of this values-investigation as they get near the end of the market research work. They are being asked for a *specific* recommendation and some of the justification for it is the inquirer's own values-in-practice, which he may perceive most imperfectly.

Notice that I am not suggesting the imposition of the *consultant's* values on the client, simply the *investigation* of the client's values by the consultant, as part of making his recommendations, *his* value judgment of *the alternatives that confront the client*.

An exactly analogous procedure is followed by the responsible psychotherapist, and the methodology of the general practitioner is—or should be—simpler only in degree. Now education, from the point of view of experimental design, is simply psychotherapy of the unsick, in most school systems, at least. The “medical model” is the proper model for educational research. We need have no fear about the fact/value distinction—it exists, but it only warns us of a difference between two phases of our professional activities. Our goal, in value matters, should be the discovery of the solution to problems of selection and rejection that require our professional expertise, and the demonstration, to those who face the problem, of the validity of our solution. Where the client does not have the expertise to comprehend the demonstration, he must, as with his doctor's recommendations, decide on its merit by using secondary indicators such as professional qualifications, the opinion of independent experts, reputation and record, etc. And to the extent that the facts are

not available or do not determine a unique solution, our obligation is to explain this situation too.

A Validation Procedure

I am presupposing a practical kind of value-reasoning here, committed to the following principles:

1. If doing something will bring about a state of affairs that people value, that is a good *prima facie* reason for doing it.

2. If there are *prima facie* reasons for doing something and none against, we should do it. (It is *not* necessary to have a *guarantee* that there is not or will not be anything wrong with our decision, any more than we demand similar guarantees before adopting and acting upon a scientific claim.)

3. If there is a conflict of supportable *prima facie* reasons, due to an interpersonal conflict of interest, appeal must be made to a general moral principal. (Other types of conflict of reason are settled as in any case of conflicting evidence.) Only one appears defensible and it is all that is necessary. It is the principle of *prima facie* equality of rights for all parties to the dispute (explained in 5).

4. This egalitarian principle can be defended on the temporizing ground that we are already committed to it—politically in a democracy, and theologically in almost all systems of religious ethics. Or it may be defended directly by a consideration of the advantages and disadvantages of this and alternative allocations of rights as solutions to a strategy problem in game theory.

5. *Prima facie* equality of consideration means actual equality of consideration except where inequalities can be defended on the basis of equality. For example, providing the President with a bodyguard is an inequality of consideration, but it does not represent an undemocratic or unjust or immoral arrangement because it can be defended in terms of the advantages *for everyone* of continuity in government, defense of the head public servant from a jeopardy which would, if serious, deter enlistment of the best men as presidential candidates, etc. It is easy enough on this basis to justify a system of law, including (a) punishment for the transgressor, (b) justice in its administration, and (c) many other values from the usual moral systems, with some appeal to certain empirical claims about the consequences of certain types of behaviour, e.g.—the absence of previously announced penalties for misdemeanours.

6. In particular, certain attitudes (values, wants, etc.) can be criticized as immoral if alternatives are humanly possible and would be more consistent with the equality axiom, i.e.—with morality. The need to eat or avoid pain can hardly be morally questionable, but a passionate desire for aggrandizement or riches at the expense of others' happiness

or legally rightful property can be so criticized. Hence, we do not accept as fixed constraints on the process of equal consideration the present values of the participants. Where conflict arises, those with morally indefensible values are accorded less than equal consideration in the distribution process. Hence, when a case can be made along these lines, we can reject certain demands as illicit, e.g.—demands for the death penalty as a kind of social vengeance.

The advantage of this approach to value-reasoning is that recommendations based on it are defensible in the same way that eating is defensible, i.e.—as means to human ends, and they provide us with a basis for social action no more and no less empirical than the basis for engineering action. The system thus developed itself encourages extreme caution in taking actions that seriously and adversely affect human beings except where the evidence is extremely strong, but this essential conservatism will also be found in the engineering field where very large investments are involved.

In short, the popularity of a value does not by itself guarantee its validity in any way. If, however, a value is held *and* if it cannot be shown to be a derived value based on faulty reasoning or false premises, or to be in conflict with other values of the same person which are more important to that person, or to be in conflict with values of others which do not themselves exhibit fatal flaws of these kinds and which are more important from the moral point of view, then acting to promote this value is justified.

The Practical Consequences for Education

I have been dealing in rather abstract terms for a few paragraphs, but the practical implications are extremely powerful. I am saying that the question of which curriculum or educational procedures a certain school system *should* adopt is simply part of the professional social scientist's task to *discover*. I do not say to *determine*, for that is a matter of where the power lies and it lies elsewhere. But the *problem* is just a problem in the applied social sciences. So far, perhaps, it may still seem that we have a relatively bland conclusion. But it follows automatically from this conclusion, as special cases follow from a general case, that it is simply a problem of applied social science to decide on the way in which communism is discussed, the kind of disciplinary standards that are imposed, the use of physical punishment to enforce these standards, the expulsion or transfer of undisciplined students, the application of skill-grading instead of age-grading, the use of programmed texts or computer-controlled learning, the introduction of merit increases and differentials, the treatment of religion as a sign of cultural backwardness, the encouragement of overt criticism of U.S. foreign policy, the American Legion, motherhood, and marriage. These issues are *in no way* properly the province of the untrained citizen, taxpayer or not. I stress again that the

political *power* may ultimately lie with the lay citizen, as it does on matters of the allocation of funds between research on cancer and on schizophrenia, but that in no ways justifies his making the decision. The opposite view, which has been foolishly fostered in this country, is unfortunately encouraged by the legal—indeed mystical—enshrinement of local control of schools, and is ludicrously identified as either a shining product or a fundamental pillar of democracy despite the striking evidence to the contrary in this and other countries.

Of course I am well aware of the extent of the disagreement between professional social scientists about many of these issues. Indeed, if it is necessary to produce a definite answer now, many of these questions might as well be settled by the citizenry (or by the flip of a coin) as by the experts because of our present lack of data. But of course the scientific position here must be that *no action* based on confidence in *either* answer is appropriate, *not* that actions based on confidence in *both* are perfectly appropriate. The rational strategy when answers aren't known is *not* to suppose that *both* answers *are* known. This confusion is as serious a consequence of the value-free position as the failure to recommend an answer where it is clearly supported.

So the lack of answers is *not* grounds for a lack of recommendations, but grounds for a different and more complex and cautious kind of recommendation, e.g.—of experimental trials of each alternative. But many of the above issues provide us with a considerably more promising prospect of definite answers. Now often, where the discussion is already well advanced, it is too technical for the average parent. It is then the job of the educational psychologist to draw the conclusions, *including* the recommendations.

Student Values

The random list of issues just given includes a number that refers directly to student values, and almost all affect them indirectly. Take two particular examples: the basic disciplinary problem of maintaining sufficient order in the classroom to make teaching *possible*, and the teaching of scientific method or “critical thinking” in the social studies areas.

The idea of public education does not merely *encourage*, it *presupposes* sufficient discipline in the classroom to enable the teacher and pupils to perform their assigned roles—and so of course it requires the imposition on the student by the teacher of a very definite behavioural value-system. And either expulsion or corporal punishment of the trouble-makers *may* have to be part of the teacher's repertoire if he or she is to discharge this fundamental obligation to the other students and the society. The justification of this kind of value-conclusion, in certain circumstances, is perfectly straightforward.

We are not obliged to call for a local referendum on whether to teach genetics and comparative psychology or sociology according to the presently best supported views, whether or not these views have consequences

which are offensive to an ardent racist majority in a given school district. If we are attacked for such teaching, we regard the attack as entirely improper. This is a moral judgment for which there are both moral and sound practical reasons. Teaching the truth, i.e.—best-supported theory, is a pretty well-based (cf. Lysenkoism). And in doing this we will concurrently teach values of two kinds, *if we teach well*. We will be teaching the general value of objectivity, of the scientific approach, as the most effective *way* of arriving at the truth. And we will be teaching how to apply this general method to socially and practically important issues. Amongst other examples, this will involve teaching that (or teaching the student to discover that) many of the segregationist arguments are simply ignorant nonsense. Of course this kind of teaching affects their values—not automatically or simply but frequently significantly. There is only one alternative to value-directed and value-affecting teaching in the social studies and related areas and that is not just cowardice but *incompetence, professional incompetence*.

The Moral Issues

Clearly a moral element is highly important in some of these value issues. Of course, I am suggesting that the *properly trained* social science specialist is better equipped than the average citizen to decide on the moral rights and wrongs of issues. Not only the effect of certain moral standards on behaviour is a proper object of study for the social scientist (e.g.—the anthropologist), but the very formulation and justification of these standards is an exercise in the applied social sciences, in the augmented contemporary form that includes game theory, comparative law, decision theory, and other methodologies. Moral standards are simply the behavioural imperatives associated with a particular kind of social institution, a morality or ethic, which is identified by the precedence status of its rules, the type of sanctions associated with infractions of them, their range of application, etc. As with other social institutions we can ask how effectively a morality serves the “purposes of the society”, i.e.—the needs and wants of its members. That is, we can *evaluate moralities*. Not just in terms of our local morality, but in terms of the basic needs which provide the foundation of all moralities.

It is true that morality is the most subtle and complex social institution of all, and it is so loaded with our own emotions and mythologies that social scientists have typically either denied the possibility of making normative judgments about it or made such judgments on the basis of excessively crude caricatures of it. Nor have the philosophers been of much assistance. Since Mill's death, the discussion of utilitarianism has been marked by a level of criticism not worthy of his contribution. Important criticisms do exist, and have to be met, and can be met along the lines indicated earlier. The key move in meeting them is the recognition that the original version of utilitarianism took too permissive a view

of existing utilities. It is inadequate to recommend the choice of actions or rules on the basis of maximizing expectations of current utilities (however *that* ambiguous recommendation is interpreted). It is essential that one regard each individual's attitudes as parameters and not constants in the assessment of behaviour, parameters that are not only functions of time but also partly functions of our own deliberate decisions. The social scientist is no stranger to assessment of the functional or nonfunctional social role of attitudes, which are (or reflect, or embody, etc.) one kind of value. And that kind of assessment is precisely and entirely what is involved in the evaluation of moral standards and hence in the moral evaluation of behaviour.

In our particular society, of course, it is widely maintained that morality is somehow the province of, or dependent upon, religion, but this view receives less and less support the higher one moves in the echelons of theologians—it is in fact a crude and indefensible view avidly propagated by lay churchmen. For, in theistic religions, morality is normally represented as the wishes or law of God, and this gives rise to two unsolvable problems: providing a satisfactory proof of the existence of God and showing that, if He did exist it would be morally obligatory to obey His wishes. That neither proof will ever be furnished is made extremely probable by two millenia of unsuccessful attempts and certain logical peculiarities of the problem.

In the absence of any such proofs it is, on the other hand, perfectly easy to demonstrate that there is strong survival value in a group morality; and possible, though not easy, to show that the morality based on the principle of equal rights is the optimal one. From the axiom of equal rights, together with various facts about the organization and institutions of the society, it is possible to derive the secondary values of justice, honesty, truth, trust and so on.* Morality, in this secular version, has a very simple formal structure and a very simple justification, and unless an alternative version can be objectively justified, we have the very best of reasons—the survival and welfare of each of us—to support, propagate, and work out the ramifications and applications of this system to practical life. Politically speaking, the advantage of the system just described is that its only axiom is a proposition which this country takes to be definitional of its form of government, and hence there can scarcely be any objection to teaching its consequences in the public schools. Where these consequences conflict with the moral views of various sects, the sects are to that extent guilty of undemocratic values and would have to demonstrate the error in democracy before they could make a case against this kind of value teaching in schools. (An obvious example of the consequences of undemocratic values is support for general legislation preventing the availability of efficient means of birth control.)

*See "Morality" in *Primary Philosophy*, Michael Scriven, McGraw-Hill, 1966.

So the basis of morality is simply a matter of the relative social efficiency of different attitudes towards the rights of others, and, little as we can say for sure about that, it is all we can say about morality. In short, we can justify particular moral judgments and thus justify teaching them where relevant in exactly the same way as we can justify teaching particular scientific assertions that follow from a general scientific theory for which there is good evidence.

Teaching Values

This kind of conclusion makes some people extremely nervous. It should not. I say that the justification of moral and scientific claims proceeds in *exactly* the same way and this has three important consequences.

1. We teach as facts only those assertions which really can be objectively established (such as the immorality of the death penalty and the possibility of justifying treason and suicide in certain cases); others we teach as hypotheses.* Hence, we do not violate the rights of others to make their own choices where choice is rationally possible, nor their right to know the truth where it is known.

2. Good teaching does not consist primarily in requiring the memorization of conclusions the teacher thinks are true, but in developing the skills needed to arrive at the test conclusions. Of course, this is especially crucial in moral matters, since conclusions without understanding of the arguments for them are rejected as soon as they conflict with inclinations. This is the distinction between teaching and brainwashing, and it can only be implemented gradually, since some values—a degree of obedience to parental commands, for example—must be indoctrinated in the infant before he can understand the reasons for them. But the distinction is absolutely fundamental because it is a moral obligation (as well as a pragmatic one) not to force on others views which they are given no chance to assess.

3. That certain conclusions should now be treated as established does not mean they cannot ever turn out to be wrong. The quantum theory and the death penalty and the use of cigarettes may have to be reassessed in the light of new evidence, but that in no way justifies tentativeness in discussing their present status, which is exceptionally clear and well-documented with respect to many (though naturally not all) of the most important questions about them.

So I provide answers of a very clear kind to the first two problems: we *can* justify teaching values (which we already do) and we *can* justify

*It is irrelevant that these issues are still controversial. The only relevant question is whether each side has an equally defensible position, in the light of all we now know. We can't make omelettes without breaking eggs and we can't make social progress without treading on somebody's toes. That's the name of the game.

particular values to teach. It does not follow from the fact that the answers to these questions are clear that we are now in possession of final answers to all *specific* questions about value. Of course, the answer that we do not now know the answer is itself a clear and extremely valuable answer to a value-question. Though it is not *as* valuable as a “final”, i.e. —a partisan and well-supported—answer would be, it is just as *clear*, and it makes certain actions indefensible and others proper and thus often provides just the information required by the inquirer.

A final warning and a final note of encouragement about value instruction: it is of course essential to distinguish cognitive and affective capacities here, and to direct educational effort along both dimensions. Moral analysis in particular and value analysis in general, are extremely complicated disciplines in which the cognitive methodology is not that of physics or mathematics or literature, but that of the law—and they must be taught for as many years as it takes to make a good criminal or constitutional lawyer, though the teaching can begin before kindergarten and be concurrent with other schooling. We have an absurd idea that an hour or two a month in optional Sunday schools will take care of this prodigious task of intellectual training. We fully deserve the level of moral discussion that results from such non-education.

Moral behaviour requires moral motivation as well as moral insight, and the mainspring for that (for an egalitarian morality) is identification with others, empathy, sympathy This, too, can be taught, from the very earliest ages, but not by parroting the results of the cognitive research. It can be taught by role-changing games, by tests of prediction skills about the behaviour of highly different others, by the use of highly graphic audio-visual material and by direct field experience supplemented with appropriate interviews and discussions.

So moral conclusions and moral behaviour should be taught and taught about if for no other reason than that it's immoral to keep students ignorant of the empirical punch behind the morality behind the law and the institutions which incorporate this country's virtues and permit its vices.

Problems of Testing

A final word or two about the methodological problems: I want to stress a consideration of scale. There isn't anything particularly difficult about thorough empirical studies of value and value-change agencies, if you get the problem on the right scale. Designing, performing, and analyzing such experiments must be seen as a major research operation, for the value structure of even a single individual is as hard to untangle as the fiscal policy or power-structure of a large corporation where overt deception, self-deception, and just plain obscurity combine to provide truly formidable obstacles. Perhaps we can perform impressive economies by ingenious experiments in the values field, perhaps we can run a decent

values inventory for one-tenth the cost of a corporation survey, but we certainly cannot perform miracles and do it for one one-thousandth or one ten-thousandth of the cost, especially not until we have run enough of the big studies to validate short-cut instruments.

When it comes to showing what factors *cause* value change on a large scale, a *minimal* cost model must be the lung cancer studies, whereas in practice our model is a marketing study on the effect of a new blue can on the sale of Hamm's beer. Sales are observable—even lung cancer is indirectly or (post-mortem) eventually observable—but *we* are dealing with a *highly abstract* inferential construct as our dependent variable. That means we must have much more elaborate confirmatory and investigative techniques. In the time dimension, heavy smoking, unlike a change of container, has to continue for a considerable time before it has any effect on the dependent variable and for this reason provides a more realistic analogue to a value-change study. Now think of the tens of millions of dollars behind the smoking study and ask yourself where we can find a hundredth of that amount behind a values study.

Realistically, in the absence of such funds, three courses of action recommend themselves:

1. A carefully argued presentation, every time we go up for funding, as to the minimal scale that is necessary for socially meaningful results, backed up by general arguments to this point formulated as statements of policy by our professional associations. (Just keep saying in a loud voice, "We *still* don't know if any kind of psychotherapy is effective, after fifty years of research on the wrong scale. Is that what we want here?")

2. Where we can't go for *socially* significant results, we *can* go for *locally* significant results. Define the research problems in terms of values which really matter in a particular small community—a village, a campus, a dormitory, a classroom—and you can afford to tap a very narrow range and still get substantial magnification of the sensitivity of your instruments and the leverage of your results via the emotional loading: you can even afford to go for explicit values only, and *still* get useful results.

3. Most fundamentally, *improve the instruments*, which are still about as appropriate as stone axes in an electronics workshop. Instrument *development* (not validation for general use) can be done with a sample of half a dozen subjects, and there are plenty of hints around as to how it should be done, in the occasional good work and in the techniques of other fields. We've had the Vernon-Allport-Lindzey work around for 15 years and yet three-quarters of the more recent tests make errors they show how to avoid.* It is a sad commentary on the state of the art that

*Providence has rewarded them with a copyright on the ideal acronym—only their work can be called *the value study*.

these tests should be selling well enough to encourage publishers to produce them. But there are a dozen unexploited avenues here. Why aren't we using Q-sort and R-sort methodology, forced and free, with all its advantages and highly developed techniques? (There is a hint of how this might be done in Carl Rogers' use of self-sorts for the "ideal self" in the Chicago Counselling Center work on nondirective therapy.) Why such crude uses of semantic differential techniques, which have been so far refined in the study of verbal behaviour and perception? Above all, why aren't we applying what is perhaps the key feature of programmed texts—a feature wholly immune to any of the criticisms of that technique and one characteristic of many of the best-validated instruments in the total test repertoire such as the Multiphasic—the "fragmentation" technique? We can't just throw a whole way of life at someone for rating, as Morris does, and hope to get the subject's value structure out of the single response. Even using 13 of them assumes that factor analysis has an independent intelligence somewhat transcending God's. Values are what determines the subject's response in a half-way decent instrument, and *the most direct way* of getting to them without hitting them over the head in the process is the best. An important and clever step in the right direction can be found in Roy Carter's study, where the subject rates on a 5-scale *each point* made by the discussants in an imaginary debate about the best way to run a society. Here we approach a factorizable situation, and with more rephrasing of the same points to check our interpretations, and a more limited scope, the instrument (or a family like it) could become extremely valuable. Finally, if we were to elect a single study as the most ingenious and conceptually significant research on values, I have little doubt that the Hartshorne and May study would win in a walk. And when was it published? 1928! The shame of it all! Here is a study which has the advantage of getting at values that are embodied in action, but values which are still rationally inferable from the actions (unlike the leap of blind faith in the *They Went to College* approach), *and* which additionally revealed the startling and important multifactor nature of the concept of dishonesty or deceitfulness. But where are the refinements and replications? Where are the extensions of this to other value concepts? Well, one can at least say that in values research there is plenty of room for good work.

Concluding Questions

I want to end with a question. Much of the empirical work raises deep and important methodological questions—for example, the Hartshorne and May work raises the important question of the criteria for utility of concepts with non-correlating components. These questions require further thought by the fraternity. But I shall ask a simpler and more practical value-judgmental question. We are all familiar with the discrepancy between explicit and implicit values, between avowed or espoused and real values. One's natural tendency is to view this with distaste,

or regret, as a sign of dishonesty, or lack of self-knowledge. But is this always the case—may not the distinction serve useful purposes, even be justifiable; in particular, should it be an educational objective to coalesce espoused and actual values? To put the question in a form that supports the other answer: should not ideals always be ahead of one? How far ahead before they are unrealistic or discouraging? The justification of almost any program of child education depends on the answers. But we know *nothing* useful about the answer. How little we know about the role of values!

There has been a resurgent interest in political decision-making in the local community and the structure of relationships among influential decision-makers. This study attempted to assess the accuracy of teachers and school administrators in perceiving who are the decision-makers and influentials in their own community.

A. KOWALSKI and W. D. KNILL

The University of Alberta

Teachers and Educational Decision Making

Introduction

New trends and developments in educational theory may lead to controversial changes. The attempt to implement some of these changes will emphasize the importance of the decision-making process at the community level. Formulating policies has often been frustrated and hampered by an inadequately informed public. If education is to advance in the direction proposed by educators, a clear understanding of the decision-making process is required. The problem of being cognizant of the social and community forces which impinge on the decision-making processes within a community appears to be of major concern to the teacher and the administrator alike.

Social power at the community level is not distributed evenly or equally but is concentrated in a limited number of power figures who, in effect, strongly influence community policies. Teachers must realize that they operate within a community framework structured by these individuals. Consequently, the following questions are of concern to educators. Who are the initiators of political action? What persons or groups influence the local political hierarchy? Which individuals constitute the community power structure?

Although these questions are but a sample of a plethora of queries, they are very important if knowledge about the community power structure and its relevance to education is to be obtained. Teachers must not only be cognizant of the process of decision-making, but even more important, they must find a means of influencing community decision-makers. It would appear that the activity of teachers in community affairs along with their social contacts are potentially relevant factors in the community power equation.

Research Design

This study was designed to investigate the degree of teacher perception of, and teacher participation in, community decision-making. To a limited degree, teachers' expectations for, and attitudes towards the community were also explored. The theory and research of political sociology was utilized in an attempt to determine community leaders and to discover the degree of interaction between teachers and school administrators with these community leaders and community activities.¹

The community under study is the cultural and commercial centre of a large river valley and is situated over 200 miles from a large metropolitan city. This small city, which offers industrial sites for prospective industries, a good labour supply, and many social, religious and recreational facilities, has educational facilities comparable to most cities.

The research was based in part on the data gathered in the summer of 1964 by R. Laskin.² The remainder of the study was based on structured interviews designed by the researcher to guide and facilitate the interviewing procedure for each respondent. The interviewing of the power leaders was begun by asking the city Secretary-Treasurer, who was assumed to be influential, "Who, in your opinion, are the persons in the community who have the most influence? Please list the names as they come to you and don't hesitate to name yourself if you feel that you should." The Secretary-Treasurer was then requested to rank the nominees on his list in order of importance of influence in general community affairs. These individuals were then interviewed, and the individuals they in turn nominated were subsequently interviewed. This process was repeated until no new names were added to those mentioned in previous interviews.³ The final list contained thirty-three names which constituted the group accepted as the power elite of the community.

The next phase of the study was developed from this project and utilized a group interview technique. The structured schedule utilized in the previous project was modified and adapted and formed the first part of the instrument. The remainder of the schedule contained items gathered from related studies and dealt with community participation and association with the community decision-makers. The schedule was administered during the staff meeting. Teachers and administrators were requested not to discuss any of the items while completing the instrument, but were free to ask the interviewer any questions which seemed relevant to the assignment.

Treatment of the Data

A descriptive analysis of the reputational data gathered in the summer of 1964 was undertaken. The identification of power figures as perceived by teachers and administrators and these influentials and their rank order were compared to the list of influentials revealed by the community influentials themselves.

The remainder of the teachers' responses, relating to certain demographic factors, and participation in community affairs were coded and recorded on IBM cards. Response distribution was tabulated and a trend analysis completed. To categorize the respondents according to the extent of activity in the community, a Community Activity Index was developed. This index was based on teachers' responses to questions centred on their participation and association in community affairs. These responses, which were assigned values, permitted the categorization of the responses as being low, medium or high in community activity.

An Accuracy of Perception Index was devised in order to assess the degree of accuracy with which teachers could identify the community power structure. This index was based on the list of community leaders revealed by the research carried out in the summer of 1964. A comparison was made of teacher nominations of community leaders with this list of names. This allowed teachers to be classified as either low, medium or high in their ability to perceive the community power leaders.

Results

Certain major questions rather than hypotheses were advanced in order to guide the research design. These questions were answered in terms of the data and permit the following statements to be made:

(1) High School teachers were able to identify the community decision-makers more accurately than Junior High or Elementary teachers.

(2) Teachers' accuracies in perceiving the community power structure are related to age, marital status, and length of residence in the community. Married teachers who are over thirty years of age and who have lived in the community for more than five years perceive the power structure most accurately.

(3) Teachers perceive their superordinates as influential in general community affairs as well as in school matters. The superintendent is perceived as having the most influence in the community followed by school board members and principals, respectively.

(4) Teachers do not perceive themselves as being influential in community affairs, and, in fact, do not believe that they possess an inordinate degree of influence even in school affairs.

(5) Teachers perceive school officials as a group of decision-makers who do what they themselves think best, while they believe that city officials do not often do what they think best but are influenced by certain community power figures.

(6) Administrators are more active in community affairs than high school teachers who in turn are more active than Junior High personnel. Elementary personnel are the most inactive category of teachers. Teachers and administrators as a group are quite inactive in community affairs, and in fact appear alienated from the community. However, it was noted

that approximately one-quarter of the teachers, the majority being male, aspired to more influence in general community affairs.

(7) Teachers' activity in the community is related to certain demographic characteristics:

(a) Teachers over thirty are more active in community affairs than teachers under thirty years of age.

(b) Male teachers are more active in community affairs than are females.

(c) Married teachers participate in community affairs to a greater extent than do single teachers.

(d) Teachers with more than two years of professional education participate in general community affairs more than teachers with less than two years.

(8) The locus of teachers' social activity is other teachers. More specifically, single teachers associate with other teachers more frequently than do their married counterparts.

(9) Single teachers associate professionally with other teachers to a greater extent than do married teachers. Married teachers associate professionally with the principal and vice-principal more frequently than do single teachers.

(10) Teachers associate professionally with their principals more frequently than they do with the vice-principal or the superintendent.

Implications

In the Canadian school system, policy-making is the function of elected representatives. If it is assumed that school programmes are shaped by elected representatives, and are a reflection of the population's desires, and if it is desirable that professional personnel in a school system be concerned with public opinion and all its attendant features, then teachers and administrators should be encouraged to participate in community activities.⁴ School personnel, by participating in community affairs, would have a more accurate perception of the community power structure. Since the community power structure makes many of the key decisions in the community, including educational decisions, contact with these individuals would seem desirable.

Although educators lack influence in a formal sense, they may make their influence felt through informal means; that is, by frequent informal contacts with community bodies who are the *vox populi*. Thus it appears that if teachers are to help influence community leaders, then principals and superintendents should encourage them to participate in community activities. School administrators, through example and by advocacy of teachers' participation in community affairs, could create a high degree of rapport and improved communication with the community at large.

Informal contacts with community leaders may prove to be one of the more effective and efficient means of obtaining the necessary resources for the implementation of an effective educational system.

REFERENCES

- ¹ A. Kowalski, "An Analysis of Teachers' Perceptions and of Participation in Community Decision-Making," Unpublished Master's Thesis, The University of Alberta, Edmonton, 1965.
- ² R. Laskin conducted a study of community leadership and influence as part of a larger study sponsored by the Alcoholism Foundation of Alberta. The procedures used to ascertain the leadership structure is further developed by Serena Phillet, "An Analysis of Community Influence—Some Conceptual and Methodological Considerations," Unpublished Master's Thesis, The University of Alberta, Edmonton, 1964.
- ³ Charles M. Bonjean, "Community Leadership: A Case Study and Conceptual Refinement," *American Journal of Sociology*, LVIII (1963), pp. 672-681.
- ⁴ A further distinction between "public" and "administrative" policy-making is developed by W. D. Knill, "Community Decision Processes: Research Strategies," *The Politics of Education in the Local Community*, (Danville: The Interstate Printers and Publishers, Inc., 1964), pp. 77-91.

In this article the author summarizes the results of a study of university admissions procedures and makes suggestions for revision of the existing system. The data on which this article is based has been published in a series of articles in this journal.

D O N A L D B. B L A C K

The University of Alberta, Edmonton

Methods of Predicting Freshman Success: Summary and Evaluation*

The purpose of this, the concluding article in a series of six that have appeared in this journal† is to consolidate the findings and speculate on the implications of a study of university admission procedures and criteria. The study examined three dimensions of admission procedures:

1. The influence of sex, faculty entered, and the size of graduating high school on the prediction of university freshman success.

2. The effectiveness of two systems of score reports: the results of Grade XII departmental examination, and the Principals' Rating scores of the same subjects tested by the Departmental Examination.

3. The relative effectiveness of five procedures using either zero-order (r) or multiple (R) correlation coefficients to predict University freshman success when:

- (a) using the individual scores from five Grade XII subject areas, English 30, Social Studies 30, and at least one score in each of Mathematics, Science and Foreign Language (R);

- (b) using the average of the six Grade XII subject scores submitted for admission (r);

- (c) using the Grade XII average and the ACE L and Q scores in a multiple regression format (R_A);

- (d) using the Grade XII average to predict Freshman average when both variables had been normalized by faculty entered (r_n);

- (e) using the Grade XII average and ACE L and Q scores in a multiple regression format to predict Freshman average when all variables had been normalized by faculty (R_{An}).

*Received May, 1965.

†Vol. XI, No. 1 (March, 1965), pp. 3-16; Vol. XI, No. 2 (June, 1965), pp. 71-82, 125-133; Vol. XI, No. 4 (December, 1965), pp. 262-271; Vol. XII, No. 1 (March, 1966).

Within the framework of the three factors of sex, size of graduating high school and faculty entered, considered in relation to the predictability of University scores, the effect of the method of predicting success, and the type of score system used to predict success, were examined with respect to the following four questions.

1. Inasmuch as the Grade XII Departmental Average is used as the basis of admission to the University of Alberta, how much more effective would admission be if the method of differential prediction was used?

2. Can the Grade XII Principal's Ratings be used effectively as a basis for admission? If so, what is the loss or gain in predictive efficiency?

3. Inasmuch as it is likely that there will be established a Canadian version of the American College Entrance Examination Board,¹ and in this event, a CEEB Scholastic Aptitude Test type instrument may become the common testing instrument, what is the effect of a similar instrument's scores on the prediction of success? Would such scores, when added to the Principal's Rating Average, improve success prediction sufficiently to equal or better the Departmental Average which is currently the basis of admission? Is the improvement sufficient to justify the administration of an academic aptitude test to all students prior to admission?

4. It is recognized that in the process of setting a specific Grade XII Average Score for admission to the University, skewed distributions result. Will the normalizing of both predictor and criterion variables improve the prediction of Freshman success sufficiently to justify implementation of the procedure before processing scores for admission?

The basic sample was the 1962 Freshman class at the University of Alberta, Edmonton, in the Faculties of Arts, Education, Engineering and Science who had graduated from an Alberta high school in June of that year. From this group, two overlapping samples were drawn. The first consisted of 1,017 cases for whom both Departmental and Principal's Rating scores were available in English 30, Social Studies 30, and at least one in each of Mathematics, Science and Foreign Language to a *minimum of six courses*. The second sample consisted of 1,010 subjects who had been tested with the ACE by the Student Counselling Services during Freshman Registration Week and for whom Departmental and Principal's Rating scores were available in six Grade XII subjects.

It was found that the size, defined by the number of teachers on staff, of the Alberta graduating high school is not as sensitive as expected. However, a difference was found between the performance and predictability of Freshman averages of graduates from large high schools (more than twenty-five teachers) and those from small high schools. The performance of the large high school graduates tends to be better than that of students from smaller schools. Similarly, sex differences were found which favoured female students generally both in level of performance

and particularly in the predictability of University marks. Faculty differences were found although it was not determined whether these were a result of the admission requirements for each faculty; the interest/apptitude structure of students selecting a certain faculty; the influence of the schools from which the students graduated; or a combination of all three. This question is pertinent in the case of the Faculty of Education which attracted more women than men, and more graduates from small high schools than students from large schools.

The following procedure was employed to answer these four questions. The validity coefficients for the sample by sex, faculty entered, and size of graduating high school from the previous articles were consolidated into Table I for the Total samples, and Tables III, V, VII and VIII for the samples from the Faculties of Arts, Education, Engineering and Science, respectively. The two lines for each sub-category of each table reflect the two different samples which were used to calculate the reported results and which were described in detail in the second and fifth articles in this series. It was found that no significant differences existed between the common validity coefficients calculated for both samples.

To judge the relative effectiveness of each method of predicting scores for the two score systems, the increase or decrease in the size of the validity coefficient over that obtained by using only the Grade XII Departmental score average was determined. No statistically significant differences were found between this score-procedure validity coefficient and any of the other nine score-procedures coefficients calculated for the same sub-sample, for the total samples, or for the sub-samples from each of the four faculties. There is some evidence to suggest that these validity coefficients are more stable than normally would be expected.² Consequently, consideration was given only to the numerical size of the validity coefficients without regard to sampling errors. To indicate the relative effectiveness of each score-procedure of predicting freshman success, the conventional Index of Predictive (forecasting) Efficiency ($E = [1 - \sqrt{1 - r^2}] 100$, where r is the validity coefficient) was calculated for each coefficient. The obtained Indices were then subtracted from the Index for the Departmental Average validity coefficient. A plus (+) sign was used to denote an increase in predictive efficiency; a minus (—) sign, a decrease. This signed value represents the percent change in efficiency. These data are presented in Table II for the Total sample, and Tables IV, VI, VII and IX for the faculties of Arts, Education, Engineering and Science, respectively.

From these tables, several general conclusions can be reached. Of all the comparisons, almost three quarters fall within the range $\pm 4\%$ of the predictive efficiency of the Grade XII Departmental average. About nine percent of the comparisons show an improvement of over 4% while about seventeen percent show a decline in predictive efficiency of greater

than 4%. The Principal's Rating Scores are, as predictors, as good as, or more effective than the Departmental scores for predicting the scores of male students from small high schools. However, they are less effective for all students in the large high schools, and for females from small high schools. With the exception of males from small high schools, the Departmental scores are generally not more than 2% more effective as predictors of University Freshman success than the Principal's Ratings.

Regarding the procedure of predicting scores, the ranking in efficiency whether for Departmental or Principal's Rating score would appear to be, differential prediction (R), Grade XII average and ACE scores (R_A), Grade XII averages (r) and with little to choose between normalized average (r_n) and Average and ACE scores (R_{An}). However, the striking finding is the relatively small differences in predictive efficiency of one procedure over another.

The use of one procedure rather than another may be considered in relation to cost. Can the additional work involved by a change from the system now used be justified if the reduction in predictive error is so small? This is a question which university administration must decide. It is the view of the author that it will be some time before the admissions pressure in Alberta will be sufficient to warrant the implementation of more complicated admission procedures. Even with automation, it is likely that the losses in time because of more complex procedures would offset the slight gains achieved over the present relatively simple procedure of asking for an average in a series of specified Grade XII subjects.

Examination of the use of Principal's Ratings is more fruitful. While generally slightly less effective than the Grade XII Departmental Average, the fact that they are available two and a half to three months before the Departmental scores offers positive time advantages for their use in academic guidance before the students leave the high school, for administrative planning in the University, and for parents and students alike in making the necessary preparations for the student to attend university. In the opinion of the author, the time factor resulting from provincial admission based on these marks would more than offset the slight loss in predictive efficiency. Although no evidence is available, it is possible that the opportunity for positive academic guidance under conditions realistic as to choice of faculty before the student leaves the school, may reduce the margin of predictive efficiency between the two score systems. The use of academic aptitude test scores coupled with the Principal's Rating average does reduce the predictive gap only slightly in many instances. However, when the Principal's Ratings are relatively poor predictors, aptitude test scores do increase the predictive efficiency appreciably. Their overall use is probably not warranted, but the fact that such test scores are normally part of a student's case history and necessary for effective guidance suggests that they be used if they are available for

all students. However, their real use, namely for admitting students with marginal academic records but with high test scores, has not been examined in any of the Alberta studies. Like many other aspects of admission to be referred to later, full information will not be available until the data from experimental admissions are examined. Costs in an operational setting again will be a primary determiner of implementation.

It is clear that the findings of this series of studies align themselves with those supporting the high school average as the best single predictor of University Freshman success. It is equally evident that the Principal's Ratings as presently reported offer considerable operational advantages with only a small loss in predictive efficiency.

One other major conclusion arises from a review of the literature and consideration of the findings of this series of studies. It would appear that little is to be gained in Alberta from continued admissions research of the kind illustrated by this series of studies. Over the past three decades little improvement, in terms of increase in the size of validity coefficients, in the prediction of success has been achieved by using achievement tests.

Gains have been made when predictor variables manifest low inter-correlations. This occurs with aptitude type batteries but is not a feature of Departmental Examination scores. Until the University is prepared to use scores from such instruments in lieu of formal achievement test batteries (viz., Departmental Examinations), and to recognize high school recommendations, little further improvement in validity coefficients can be expected.

The effectiveness of Departmental scores seems to hinge on two points: first, the tests are closely articulated with the University program; and second, of those attempting the Departmentals, only a select proportion are admitted to the University. It could be hypothesized that if we prepared achievement examinations solely for the purposes of University admissions little change would occur in the validity coefficients from those presently obtained using the Departmental examinations scores. There is much to recommend separate testing procedures on other grounds. For example, such a testing program would free the high school to attain those objectives it claims to seek for all its pupils. Undoubtedly, there is much that can be done to improve the present instruments, chiefly through the adoption of testing protocols already proven valuable, and through a shorter testing period per instrument than now exists. It would seem inevitable that such developments can best be promoted on a national level although in Alberta good progress is already being made in this direction with the Departmental examinations. Even with such specialized instruments, their use will need careful re-examination. Undoubtedly, their predictive efficiencies, in the event of a national battery, would fall to a level close to those reported in the literature for similar

programs. Such scores would have to be supplemented by Principal's evaluations and some provision made for providing performance data back to the schools as a basis for their validation of these ratings. A testing program specifically aimed at University admission but offered earlier in the senior year than the present Departmentals offers the most obvious solution and is the one which has received widespread acceptance in the United States. However, it is not likely that validity coefficients will increase over those for instruments presently in existence. The advantage of such a development would probably be achieved in the operational ease of developing and conducting the testing program *per se*, and freeing the curriculum of the high schools from the present constraints.

There seems to be little evidence that non-intellectual predictor variables offer much hope for improvement. Viewing the voluminous research on motivational measures, interest, persistence, achievement motivation, study habits and attitudes, and biographical and personal histories, the findings have been inconsistent and generally offer little improvement in predictive measures over those provided by the conventional measures of achievement and aptitude. It may well be that a major breakthrough will occur in this area but at this time, with the instruments presently available, this is not considered likely. But, one cannot ignore the appeal of Getzels³ in his case for admitting students with "divergent" intellectual and social interests. He claims the weakness with the above mentioned methods is that they, like academic tests, tap only those of "convergent" intellectual ability and we may be losing those who, in the long run, will prove to be more creative. The studies of creativity do offer the best hope of producing instruments to identify these students who typify the kind of graduate each institution aspires to produce.

There has been one glaring gap in admissions research at the University of Alberta. No studies of experimental admissions of students with qualifications at marked variance to those normally required have been undertaken. It is contended that little is to be gained by further descriptive studies like those of the past decade and typified by this series of studies. This is considered a matter of immediate and primary importance for two reasons: first, the high schools are developing away from the single matriculation stream that characterized them for so long. There can be little doubt that within a decade the University must be prepared to take a stand on a multi-program admission policy. Every consideration of this age of automation which we are entering points inevitably in this direction. A Canadian replicaton of the Eight Year Study⁴ is clearly in order, for Canadian universities have too long ignored the important implications of this study. This had been due primarily to the past one-program orientation of our high schools. Second, and this is already obvious, the universities are coming under increasing pressure to admit more students to facilities already becoming overtaxed. There is

danger that unless the problem of multi-stream admission is acted upon now, the flood of coming university enrollments will offer the University an excellent excuse for a laissez-faire attitude towards its admission policies. Yet, if the University is to meet its obligations to the community at large, it must develop methods to recognize those high school students who can profit from a university education *regardless of the high school program taken*. To do this will mean a marked change in present admission policies. It is argued that these changes must be considered now while the University still has the opportunity to examine experimentally a variety of possible admission policies. Already there is some evidence from one Alberta junior college that our present admission policies warrant careful experimental consideration. It reports that a third of its students who were not eligible for admission to the University succeed in University accredited programs. Many will argue that these are problems solely for the junior colleges and recommend that the University "let them worry about these things." Admittedly, these are not just problems for the University alone but for all higher education. The University of Alberta, however, is in an ideal position in Canada to display enlightened leadership by using its qualified staff and facilities to conduct pioneer Canadian research in the study of experimental admission procedures on a variety of fronts by involving the junior college as well. Without this experimental approach, in the pessimistic opinion of the writer, little is to be gained from further descriptive studies of University admission practices. Regrettably, the time left to take action before the pressure of enrollments makes experimental considerations almost impossible is disturbingly short.

REFERENCES

- ¹ D. B. Black, "Some comments on the types of problems inherent in a national testing program for university admissions", *Alberta Journal of Educational Research*, Vol. IX, No. 1, 1963, pp. 3-12.
- ² "Validity of regression equations after six years to predict freshman success in engineering", *Alberta Journal of Educational Research*, Vol. X, No. 2, 1964, pp. 125-136.
- ³ J. W. Getzels, "Non - I.Q. Intellectual and other factors in college admission", in *The Coming Crisis in the Selection of Students for College Entrance*, K. E. Anderson, editor, (Washington: American Educational Research Association/ National Education Association, 1960), pp. 23-28.
- ⁴ Dean Chamberlain, et al., *Did They Succeed in College* (New York: Harper and Brothers, 1942), p. 291.

TABLE I

VALIDITY COEFFICIENTS FOR PREDICTING UNIVERSITY FRESHMAN SUCCESS BY VARIOUS TECHNIQUES BY SEX AND SIZE OF GRADUATING HIGH SCHOOL USING GRADE XII DEPARTMENTAL AND PRINCIPALS' RATING SCORES

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Small	male*	.555	.561	.537	.549	.546	.542	.486	.543	.504	.547
	**			.531	.540						
	female	.610	.584	.580	.537	.522	.525	.551	.552	.570	.552
Large	Total	.560	.558	.545	.537	.547	.538	.518	.533	.532	.544
				.530	.525						
	male	.718	.664	.728	.658	.732	.656	.701	.615	.703	.617
Total	female	.791	.764	.730	.654	.786	.768	.730	.698	.736	.711
	Total	.735	.697	.791	.759	.746	.692	.713	.649	.713	.650
				.782	.761						
Total	male	.635	.586	.640	.599	.640	.594	.634	.594	.640	.594
	female	.679	.648	.634	.594	.679	.644	.674	.628	.679	.644
	Total	.636	.605	.683	.628	.647	.606	.642	.603	.647	.606
				.674	.609						
				.642	.603						

*based on the first sample of 1,017 cases as defined in the second article of this series.

**based on the second sample of 1,010 cases as defined in the fifth article of this series.

TABLE II
INCREASE OR DECREASE IN PERCENT IN PREDICTIVE EFFICIENCY OVER GRADE XII DEPARTMENTAL SCORE AVERAGE
BY USING DIFFERENT METHODS OF PREDICTING FRESHMAN SUCCESS IN FOUR FACULTIES COMBINED

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Small	male	+1.18	+1.58	—	+ .57	+	.96	—2.66	+	—1.63	+1.03
	female	+2.22	+ .29	—	+ .06	+	.00	+1.84	.00	+3.13	+1.91
	Total	+	.86	—	— .31	+	+1.09	— .74	+	+	.89
Large	male	—1.04	—6.21	—	—7.31	+	.21	—2.98	—10.51	—2.78	—10.36
	female	.00	—3.34	—	—2.54	+	.51	—6.01	—9.28	—5.37	—7.99
	Total	—1.41	—5.31	—	—5.70	+	.00	—3.53	—9.49	—3.53	—9.40
Total	male	— .41	—4.19	—	—3.12	+	.49	.00	—3.12	+	.49
	female	— .37	—3.12	—	—3.95	+	.46	.00	—3.95	+	.46
	Total	—1.09	—3.54	—	—3.10	+	.42	.00	—3.10	+	.42

TABLE III
VALIDITY COEFFICIENTS FOR PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF ARTS BY VARIOUS TECHNIQUES
AND BY SEX AND SIZE OF GRADUATING HIGH SCHOOL USING GRADE XII DEPARTMENTAL AND PRINCIPALS' RATING
SCORES

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Small	male	.444	.484	.389	.405	.491	.449	.401	.483	.480	.504
	female	.693	.635	.382	.401	.620	.585	.569	.553	.575	.553
	Total	.546	.553	.614	.583	.538	.504	.471	.511	.506	.519
Large	male	.733	.528	.649	.509	.655	.518	.673	.521	.674	.527
	female	.826	.769	.652	.511	.820	.765	.814	.723	.815	.732
	Total	.783	.671	.824	.758	.755	.659	.733	.616	.739	.621
Total	male	.558	.483	.534	.449	.546	.452	.555	.490	.566	.492
	female	.746	.696	.758	.694	.749	.690	.725	.658	.726	.659
	Total	.653	.607	.748	.690	.656	.583	.631	.565	.635	.567

TABLE IV
INCREASE OR DECREASE IN PREDICTIVE EFFICIENCY OVER GRADE XII DEPARTMENTAL AVERAGE BY USING DIFFERENT METHODS OF PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF ARTS

Method of Predicting Freshman Success												
Size of School	Sex	R		r		R _A		r _n		R _{An}		
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	
Small	male	+2.52	+4.61	—	+ .81	+5.30	+3.07	+ .81	+4.86	+4.69	+6.05	
	female	+3.73	—1.43	—	—2.32	+ .47	—2.17	—3.30	—4.24	—2.88	—4.39	
	Total	+1.08	+1.54	—	— .39	+2.88	— .80	—1.04	+1.21	+ .92	+1.69	
Large	male	+8.06	—8.84	—	—10.14	+ .26	—9.72	+1.86	—9.54	+1.95	—9.17	
	female	+ .29	—7.26	—	—7.53	.00	—7.16	— .85	—11.84	— .71	—10.89	
	Total	+3.37	—8.58	—	—9.41	+ .23	—9.41	—1.93	—12.97	—1.80	—12.58	
Total	male	+1.57	—3.01	—	—4.56	+ .96	—4.46	+1.56	—2.43	+2.30	—2.32	
	female	—1.36	—6.57	—	—6.01	+ .11	—6.01	—2.50	—8.93	—1.40	—8.84	
	Total	— .61	—4.34	—	—5.40	+ .51	—5.26	—1.59	—6.52	—1.26	—6.38	

TABLE V
VALIDITY COEFFICIENTS FOR PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF EDUCATION BY VARIOUS TECHNIQUES AND BY SEX AND SIZE OF GRADUATING HIGH SCHOOL USING GRADE XII DEPARTMENTAL AND PRINCIPALS' RATING SCORES

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Small	male	.495	.556	.357	.479	.393	.480	.359	.472	.385	.489
	female	.635	.607	.365	.465	.589	.568	.546	.518	.582	.575
	Total	.586	.579	.565	.514	.547	.552	.492	.504	.521	.546
Large	male	.640	.578	.565	.438	.446	.426	.441	.426	.484	.446
	female	.651	.653	.426	.420	.636	.638	.611	.606	.632	.639
	Total	.645	.631	.607	.607	.606	.571	.563	.503	.587	.538
Total	male	.470	.521	.393	.435	.394	.416	.375	.394	.397	.414
	female	.634	.607	.376	.403	.599	.580	.564	.537	.593	.586
	Total	.596	.584	.590	.535	.558	.548	.511	.497	.535	.535

TABLE VI
INCREASE OR DECREASE IN PREDICTIVE EFFICIENCY OVER GRADE XII DEPARTMENTAL AVERAGE BY USING DIFFER-
ENT METHODS OF PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF EDUCATION

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Small	male	+6.52	+10.29	—	+4.57	+1.15	+5.37	— .23	+4.94	+ .81	+5.87
	female	+5.12	+2.90	—	—3.27	+1.70	+ .21	—1.27	—3.03	+1.19	+ .79
	Total	+4.26	+3.76	—	— .54	+1.77	+2.10	—1.58	— .89	+ .12	+1.70
Large	male	+5.67	+ .91	—	— .28	+ .97	.00	+ .72	.00	+2.96	+ .97
	female	+ .33	+1.01	—	.00	+2.30	+2.47	+ .31	— .08	+1.97	+2.55
	Total	+1.08	— .08	—	—2.92	+1.70	— .84	—1.40	—5.18	+ .29	—3.04
Total	male	+3.68	+6.59	—	+1.14	+ .75	+1.72	— .04	+ .65	+ .88	+1.63
	female	+3.41	+1.27	—	—2.95	+1.46	+ .07	—1.05	—2.83	+1.01	+ .50
	Total	+3.02	+2.15	—	—1.54	+1.38	+ .71	—1.60	—2.41	— .12	— .12

TABLE VII
VALIDITY COEFFICIENTS FOR PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF ENGINEERING BY VARIOUS TECHNIQUES USING GRADE XII DEPARTMENTAL AND PRINCIPALS' RATING SCORES AND THE INCREASE OR DECREASE IN PREDICTIVE EFFICIENCY OF THESE TECHNIQUES OVER THE GRADE XII DEPARTMENTAL AVERAGE

Size of School	Sex	Method of Predicting Freshman Success									
		R		r		R _A		r _n		R _{An}	
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.
Validity Coefficients	Small	.732	.588	.609 .601	.532 .528	.625	.569	.625	.574	.625	.609
	Large	.831	.740	.730 .730	.581 .581	.776	.655	.746	.632	.769	.681
	Total	.750	.633	.661 .659	.581 .545	.691	.602	.679	.591	.702	.634
Predictive Efficiency	Small	+11.19	-1.57	—	-5.00	-1.86	-2.31	+1.86	-1.97	+1.86	+ .60
	Large	+12.71	+1.08	—	-13.05	+5.27	-7.22	+1.75	-9.16	+4.42	-4.89
	Total	+ 8.90	-2.37	—	-8.63	+2.92	-4.64	+1.80	-5.46	+3.99	-2.12

TABLE VIII
VALIDITY COEFFICIENTS FOR PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF SCIENCE BY VARIOUS TECHNIQUES AND BY SEX AND SIZE OF GRADUATING HIGH SCHOOL USING GRADE XII DEPARTMENTAL AND PRINCIPALS' RATING SCORES

Size of School	Sex	Method of Predicting Freshman Success											
		R		r		R _A		r _n		R _{An}			
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.		P.R.
Small	male	.632	.693	.623	.652	.620	.647	.576	.625	.591	.630		
	female	.621	.671	.616	.645								
	Total	.625	.671	.670	.633	.682	.659	.696	.638	.710	.647		
Large	male			.603	.644	.613	.645	.590	.623	.597	.625		
	female	.799	.788	.795	.785	.807	.788	.769	.768	.773	.770		
	Total	.909	.926	.908	.788	.908	.928	.876	.882	.883	.882		
Total	male	.818	.814	.908	.928	.816	.816	.792	.794	.794	.795		
	female			.817	.813								
	Total			.816	.816								
	male	.722	.698	.727	.728	.733	.729	.695	.709	.700	.710		
	female	.796	.800	.724	.728	.807	.831	.801	.805	.814	.807		
	Total	.737	.713	.792	.819	.746	.747	.717	.728	.718	.729		

TABLE IX
INCREASE OR DECREASE IN PREDICTIVE EFFICIENCY OVER GRADE XII DEPARTMENTAL AVERAGE BY USING DIFFER-
ENT METHODS OF PREDICTING FRESHMAN SUCCESS IN THE FACULTY OF SCIENCE

Method of Predicting Freshman Success												
Size of School	Sex	R		r		R _A		r _n		R _{An}		
		Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	Dept.	P.R.	
Small	male	+	.72	—	+3.50	+1.46	+3.46	—1.82	+1.86	—	.75	
	female	+	.39	—	— .97	+1.11	— .97	+2.44	—2.76	+	3.82	
	Total	+	.32	—	+3.27	+	.76	— .97	+1.55	—	.45	
Large	male	+	.53	—	—2.11	+	.41	—4.46	—4.58	—	3.98	
	female	+	.22	—	+4.64	.00	+4.64	—6.33	—5.22	—	5.04	
	Total	+	.14	—	.00	.00	.00	—3.24	—2.98	—	2.98	
Total	male	—	.53	—	+	.42	+	.53	—2.92	—1.54	—2.43	
	female	+	.52	—	+3.55	+	.28	— .54	.00	+	1.24	
	Total	—	.55	—	+	.45	+	.45	—2.67	—1.52	—2.56	

The assumption that different personality characteristics exist between good elementary teachers and good secondary teachers is often made. If this is valid, efficient counselling and guidance of students in teacher education programmes necessitates the delineation and measurement of the relevant traits. The authors of this study report a preliminary investigation of these differences and indicates some areas where further investigation is necessary and is being undertaken.

O. H. BOWN, F. F. FULLER
and H. G. RICHEK

University of Texas

Differentiating Prospective Elementary and Secondary School Teachers*

Introduction

The Myers-Briggs Type Indicator is intended for idiographic use in a wide variety of situations including academic and vocational counselling. It has been used periodically by the two senior authors (both counselling psychologists) in the screening of students at the College of Education, The University of Texas, who aspire to be teachers at either the elementary or secondary school level.

In the course of adducing quantitative data in support of our ever-strengthening conviction that the *sine qua non* characteristics of an effective elementary school teacher are (1) warmth, (2) a helping orientation and (3) a capacity for identification¹ we utilized scores from the Myers-Briggs to compare future elementary school teachers with their secondary peers. From their literary definitions, at least one of the Jungian typologies would appear to be conceptually related to our triad of necessary (if not sufficient) elementary school teacher qualities. The test manual² contains information on frequency and mean strength of the type preferences for female school teachers but does not differentiate elementary from secondary school teachers.

*The research reported herein was performed pursuant to a contract with the United States Department of Health, Education and Welfare, Office of Education, under the provisions of the Cooperative Research Program.

Hypotheses

Except in one regard, the rationale for employing the Myers-Briggs Inventory in this study did not subsume any a priori notions that Jung's psychological types have substantial correspondence with potential teaching effectiveness and/or teaching styles. The so-called "extroverted attitude"³ type as explicated by Jung was assumed to be more closely related to two of the three essential characteristics for elementary school teachers (i.e., *capacity for identification and warmth*) than is its polar opposite—the introverted attitude type. The primary hypothesis to be investigated, therefore, was that the prospective elementary school teachers would differ significantly from secondary school teachers on measures of extroversion/introversion, i.e., elementary school teachers would be more extroverted, less introverted (Hypothesis I).

Utilizing the definitions of the function types from the Myers-Briggs Manual, we investigated three additional hypotheses with regard to our two groups. These are expressed in null form since as far as is known there exists neither theoretical nor empirical justification for directional hypotheses.

Hypothesis II: there is no difference between future elementary school teachers and future secondary school teachers on measures of sensation/intuition.

Hypothesis III: there is no difference between elementary school and secondary school teacher aspirants on measures of thinking/feeling.

Hypothesis IV: there is no difference between future elementary school teachers and secondary school teachers on measures of judgement/perception.

Method

For a period of approximately one academic year, students (female) who were participants in the Experimental Teaching Program of the Personality, Teacher Education and Teaching Behaviour Project were administered the Myers-Briggs Type Indicator at the point of admission to the experimental or control groups.

The scored protocols of 149 students (the total number of participants in the experimental program during the year) were submitted to single classification Analyses of Variance with regard to differences in the group means on measures of four Jungian concepts: Extroversion/Introversion; Sensing/Intuition; Thinking/Feeling; Judgement/Perception. Students were divided into two groups: future elementary school teachers (N=71) and prospective secondary school teachers (N=78). Students in both groups were approximately the same ages. In view of the small N involved, no information about teaching fields was sought in connection with this analysis. For an hypothesis to receive confirmation, it was necessary that the differences in group means be statistically significant

at the .05 level on three measures of the construct under investigation, e.g., scores on the extroversion scale, its obverse, the introversion scale and the extroversion/introversion continuum or preference scores.

Results and Discussion

TABLE I
ONE-WAY ANALYSIS OF VARIANCE FOR MEANS OF ELEMENTARY AND
SECONDARY EDUCATION MAJORS FROM MYERS-BRIGGS TYPE
INDICATOR SCORES

Scale	Means		F-Ratio	P
	Elementary (N=71)	Secondary (N=78)		
Extroversion	17.03	15.10	3.69	.06
Sensing	12.41	11.26	.82	.37
Thinking	6.61	7.38	.77	.38
Judging	15.93	15.44	.23	.63
Introversion	9.07	10.71	2.59	.11
Intuition	10.86	12.74	3.38	.07
Feeling	13.08	12.96	.02	.89
Perceiving	10.34	11.28	.82	.37
Extroversion/Introversion	85.08	92.21	3.25	.07
Sensing/Intuition	97.90	103.97	1.85	.18
Thinking/Feeling	113.96	112.15	.29	.59
Judging/Perceiving	89.82	92.69	.50	.48

It will be seen that only with regard to our primary hypothesis were results close to the “traditional” level of significance. On the three measures of extroversion/introversion, means of future elementary school teachers were, in all instances, different from the secondary group in the hypothesized directions but not to a statistically significant extent. The hypothesis, while not receiving unequivocal confirmation by our data, does seem to warrant further investigation.

No significant findings emerged on the sensation/intuition; thinking/feeling; judging/perceiving indices. It will be noted that prospective elementary school teachers are somewhat less “intuitive” than the secondary group ($p=.08$) but the group differences do not approach significance on either the obverse characteristic (sensation) or the sensation/intuition continuum.

The particular subject matter field of future secondary school teachers was not taken into account in this study since the small N precluded such a procedure. This limitation of the design may have affected the

results on both the extroversion/introversion and sensation/intuition dimensions. Future secondary school teachers may differ in these characteristics in accordance with the nature of their teaching specialties, i.e., physical sciences, life sciences, social sciences, belles-lettres, fine arts, etc. These questions are, of course, susceptible to resolution by research with larger samples which is currently in progress.

Summary

Previous research⁴ has supported the proposition that desirable characteristics of future elementary school teachers are (1) warmth, (2) helping orientation and (3) capacity for identification. Two of these three (1 and 3) were assumed by the authors to be related to Jung's "extroverted attitude type". It was hypothesized that future elementary school teachers would, therefore, score higher on the Myers-Briggs measures of extroversion and lower on the measures of introversion than their secondary school counterparts. Neither this hypothesis nor three others formulated with regard to Jung's "function types" were confirmed by the data of this study which consisted of Myers-Briggs Type Indicator scores on 149 female Education majors at The University of Texas. Further investigation, however, of the extroversion/introversion dimension in prospective teacher groups appears warranted. Such research will utilize larger samples, thus making possible statistical control of the subject matter teaching field. Lack of this information may have obscured the results in this study, particularly regarding extroversion/introversion.

REFERENCES

- ¹ L. W. Sontag, and J. Kagan, "The Emergence of Intellectual Achievement Motives", *American Journal of Orthopsychiatry*, Vol. XXXIII, 1963, pp. 532-535.
- ² Isabel B. Myers, *The Myers-Briggs Type Indicator* (Princeton, Educational Testing Service, 1962).
- ³ C. G. Jung, *Psychological Types* (New York: Harcourt Brace, 1923).
- ⁴ Sontag and Kagan, *op. cit.*

In this article the author uses a simple experimental design to examine the effectiveness of three methods of developing reading ability in beginning reading.... The findings provide some support for the belief that effective oral reading to beginners produces beneficial results.

M A R Y H. N E V I L L E

University of Alberta, Calgary

Methods of Teaching Reading To Beginners*

Introduction

When a child first begins to read, he must pay close attention to individual words and to their constituents in order to identify and later recognize them. Yet we think or understand in groups of words.¹ How, then, does the beginning reader learn to relate singly identified words into the thought of a sentence?

Linguists believe that in speech intonation patterns are a powerful means of showing the connectedness and completeness of groups of words.² Recent psychological theory concerning the acquisition of grammatical structure also suggests that intonation may aid this complex language learning.³ Lloyd⁴ from the linguistic, and Gleissman⁵ from the psychological viewpoint, have both hypothesized that the intonation patterns of speech may, in reading, help to unite singly identified words into units of thought. These hypotheses are supported by the relationship shown between reading fluency, and the ability to apply normal intonation patterns to reading, and reading comprehension.⁶ Assuming that intonation does help the reader relate the words of a sentence into "thought units", how can the beginning reader learn the ability to apply intonation patterns to his reading?

Skinner⁷ and Mowrer⁸ believe that when children learn to speak they do so by means of imitation of the speech of those around them. As they are reinforced for speech which comes ever closer to an acceptable model, they also become very accomplished in echoic behaviour. This accomplishment is then used to short-circuit the slower process of progressive approximation used in the earlier language learning. Eventually, not only

*This paper is based on the author's M.Ed. thesis presented to the University of Alberta, Calgary. The Author would like to express her appreciation for the guidance received from her advisor, Dr. E. M. King.

does speech itself acquire secondary reinforcing properties, but so does the accurate echoing of speech. Although Skinner states that in true oral reading (or "texting") no echoic behaviour occurs, the learning principle of imitation is used in teaching reading by the "look and say" method.⁹ The visual stimulus (a word) is shown and the teacher says its oral form. The child may then make an overt, echoic response or he may make a silent or covert response.

Arguments in favour of the covert response state that it is faster than the overt, oral response and that it discourages the habit of vocalization.¹⁰ Skinner believes that the oral response is essential in beginning reading. Without it extrinsic reinforcement is not possible. Furthermore, the secondary reinforcement properties of speech are not utilized and neither can the child so readily respond, as a listener, to his own self-stimulation. The association of meaning with the written word is also facilitated by the oral response.¹¹ By saying the word, the mediational processes of the oral form become more easily associated with the visual symbols. Some recent studies¹² do indicate that the oral response and oral reading aid learning to read and also improve comprehension. When the child learns to read connected material, possibly the same methods that were used to teach him words could be used to teach him to apply intonation patterns to sentences.

In reading larger language structures, whose individual words have already been taught, if the beginning reader makes an oral response to the visual stimuli he will be reinforced by hearing himself pronounce words. Due to self-correction tendencies¹³ (perhaps the result of much echoic behaviour) he will try to make the words sound like normal speech with normal intonation patterns—especially if he is instructed to do this. The oral response may have the further advantage of promoting the association of the mediation processes of the spoken word with the visual symbols—improvement in this association also encourages cue reduction and a decrease in the incidence of vocalization.

Skinner states that the model for an echoic response can vary in size from phonemes to groups of words or sentences. If a model for the reading was supplied the oral response would become an echoic response which, by short-circuiting the process of progressive approximation, would be a more efficient means of teaching oral reading with the intonation patterns of speech.

The ability to apply intonation patterns to reading should not only improve comprehension, it should also help word identification and recognition during reading. When an unknown or forgotten word is encountered in context it can often be guessed successfully.¹⁴ The child, in listening to speech, makes efficient use of context to help him identify words he did not hear;¹⁵ thus the closer the written context approaches familiar speech the more easily will the reader guess a word correctly.

The word is then not a chance entity but a missing part of a meaningful language structure.

The problem is, assuming that applying the intonation patterns of speech to reading helps the child grasp sentence meaning and identify words from context, will the beginning reader's ability to apply intonation patterns be improved by an oral response and by giving a model for that response? The purpose of the present study was to carry out a short exploratory investigation of this problem to determine whether it should be studied more comprehensively over a longer period. The specific hypothesis was: reading silently, reading orally, and listening to a competent oral reading of instructional material will have differential effects in beginning reading on accuracy of word identification and on comprehension.

Method

The investigation was designed so that silent reading, oral reading and listening to the reading material formed the manipulation or treatments of an independent variable, practice before reading orally, whose effects were measured on two dependent variables, accuracy of word identification and comprehension.

Subjects

A sample of 93 grade one pupils was chosen from two schools in a new housing subdivision where the socio-economic background may be described as ranging from " . . . artisan up to professional with a definite weighting in the middle class."¹⁶ To attempt to control some variables known to affect the dependent variable, the sample met certain criteria. All the subjects were from classes where the teachers had had at least two years' experience. No subject was repeating grade one and all were from English-speaking homes. The distribution of IQ in the sample (as measured by the *Detroit Beginners First-Grade Intelligence Test*, 1937) was the same as in the limited two-school population of 183 from which it was drawn.

Over the range of IQ's in the sample of 93, the subjects were arranged in threes or triads so that each of the three had approximately the same IQ; the triads were alternately all male and all female. The subjects in each triad were then assigned randomly to the three treatments (silent reading, oral reading, and listening). Thus three groups, each of 31, were set up matched with regard to intelligence and sex. This design is shown in Table I.

Materials

The materials consisted of a test passage of 102 words based on a story with a good plot from the *Royal Road Readers*. There were no

pictures since the words carried the story. A few vocabulary substitutions were made and some sentence structures modified in accordance with Strickland's¹⁷ data on sentence structures used commonly by grade one children in their speech. The passage contained four words which even the most advanced subjects had not been taught in the basic reader lessons and seventy words which all subjects, even the slowest, had been taught. Eight comprehension questions about the test passage were also prepared, four being direct recall-type questions, and four inference-type questions which required some interpretation of the facts. Model answers to these questions, based on answers given during a pilot test, were also compiled.

TABLE I
THE SAMPLE AFTER ASSIGNMENT TO TREATMENTS

Sex	IQ	Frequency	Subjects in each treatment (by triad number)		
			Silent	Oral	Listening
M	141+	3	1	1	1
F	133-140	12	2	2	2
M			3	3	3
F	128-132		4	4	4
M			5	5	5
F	126-127	30	6	6	6
M	124-125		7	7	7
F			8	8	8
M	122-123		9	9	9
F			10	10	10
M	120-121		11	11	11
F			12	12	12
M	117-119		13	13	13
F			14	14	14
M			15	15	15
F	116	3	16	16	16
M	114-115	30	17	17	17
F			18	18	18
M	112-113		19	19	19
F			20	20	20
M	109-111		21	21	21
F			22	22	22
M	107-108		23	23	23
F			24	24	24
M	105-106		25	25	25
F			26	26	26
M	99-104	12	27	27	27
F	92- 98		28	28	28
M			29	29	29
F			30	30	30
M	91—	3	31	31	31

Procedure

During the middle two weeks of February, 1965, the subjects were tested individually. They were taken by the experimenter in rotating order of treatments from the classroom to a testing room. Subjects to be

given the silent reading treatment were handed the test passage and asked, "Will you read this story silently? Read with your eyes only and do not use your lips." No words were identified for the subject during this period but he was observed for signs of vocalization. Subjects given the oral reading treatment were handed the test passage and asked, "Will you read the story aloud? Read it aloud just as you would talk." The subject was given no help with unknown words and the experimenter tried to give the appearance of not listening. In the listening treatment, the experimenter said, "First I am going to read the story. Will you listen to the story carefully?" The experimenter, using natural intonation, then read the story to the subject.

After exposure to any one of the three treatments the procedure was the same for each subject. To measure accuracy of word identification the subject was asked to read the story aloud and, as he did so, his errors were marked on the test passage printed on the individual record sheet. At this reading, unknown words were told to the subject and recorded as errors. Finally this second reading of the passage was rated fluent or non-fluent by the experimenter using as a criterion Harris' definition of fluency, "... reading proceeds smoothly and rhythmically. The words are grouped in phrases, and meaningful thought units are indicated by appropriate pauses and inflections of the voice."¹⁸

The passage was then taken from the subject and to measure comprehension, the experimenter asked the subject the comprehension questions orally. The subject answered orally and the experimenter recorded the answers *verbatim* on the individual record sheet. The number of errors made during the oral reading constituted the word-identification error score. The comprehension questions were scored by comparison with the model answers.

Results

To compare the effects of the three treatments on accuracy of word identification and comprehension, analysis of variance was used.

For the accuracy test the difference between the means of the three groups was not significant. However, when the means for the errors in identifying the four "new" words were compared, the difference between the means was significant at the .001 level. The related t-test showed that the differences between the listening and silent reading means and the listening and oral reading means were significant at the .001 level whereas the silent oral reading means were not significantly different.

For the comprehension test, the differences between the means of the three groups was statistically significant ($P < .01$). A series of two sample t-tests showed that the differences between the listening and silent group means and the listening and oral group means were significant at the

.001 level. There was, however, no statistically significant difference between the means of the oral and silent groups.

Since the members of each triad had been matched before assignment to treatments, results could be ranked within triads and compared by means of a χ^2 test. The expected frequency for any rank for any treatment was 10.3 and therefore if observed frequencies did not differ significantly from this the ranking could be assumed to be random and the treatment assumed to have had no effect. As can be seen from Table II, in the accuracy test the listening subjects were ranked first a significantly greater number of times than the other two subjects and the oral subjects first a significantly fewer number of times ($P < .05$). For the comprehension test, the same pattern occurred but significance for the listening group was at the .01 level.

TABLE II
FREQUENCY OF RANKS FOR SILENT, ORAL, AND LISTENING
TREATMENTS FOR ACCURACY AND COMPREHENSION
TESTS

Rank	Silent		Oral		Listening	
	Accuracy	Comprehen- sion	Accuracy	Comprehen- sion	Accuracy	Comprehen- sion
I	10	4.5	4*	7.5	17*	19**
II	10	16.5	12	7.0	9	7.5
III	11	10.0	15	16.5*	5	4.5
N	31	31	31	31	31	31

** $p \leq .01$ (from χ^2 test)

* $p \leq .05$ (from χ^2 test)

The relationships between the two tests of accuracy and comprehension within each treatment were tested by means of Pearson product-moment correlation coefficients. The correlations are given in Table III.

TABLE III
CORRELATION COEFFICIENTS FOR ACCURACY AND COMPREHENSION
TESTS AND SUBTESTS WITHIN SILENT, ORAL, AND LISTENING
TREATMENTS

Tests	Silent	Oral	Listening
Total Accuracy—Total Comp.	0.68**	0.11	0.51**
Total Accuracy—Recall	0.62**	0.02	0.53**
Total Accuracy—Inference	0.51**	0.15	0.29

** $p \leq .01 : \alpha = 0.45$.

Of the total 93 subjects, two (one “oral” and one “listening”) were rated fluent readers.

Discussion

Those subjects who read silently were for both accuracy of word identification and comprehension, not significantly better than those who read orally or those who listened to the passage before reading it but for the silent group comprehension was related to word identification. Sixteen of the 31 subjects in the silent group were found to be vocalizing although told specifically not to use their lips, a proportion in agreement with the results of Sheldon and Hatch.¹⁹

In the oral group comprehension is shown to be independent of the number of reading errors. McCracken, who found the same results, suggests that this may be because the poor readers are told many words by the examiner²⁰ but, in this investigation, the examiner told words to all three groups during the test reading. However, the presence of a strange examiner and a reading method so unlike that usually followed in the basic readers may have disturbed some of the better readers and caused them to word call. Although this lack of relationship and rather poor general performance does not agree with the results obtained by McNeil and Keisler, and Brinkley *et al.*,²¹ the means for the oral group for both accuracy and comprehension were not significantly different from the means for the silent group.

The listening treatment did not cause a significant overall decrease in error scores for this group but listeners performed better than the other two subjects in each matched triad and correctly identified the four "new" words much more successfully than those subjects given the silent or oral treatments. The listening group, hearing in the practice period the exact phrases they would read, were often able to identify important key phrases or words during the reading. In the comprehension test the listening group performed significantly better than the other two groups. Although from Table III it could be argued that in the inference questions the high scores may have reflected listening comprehension *per se* and not reading comprehension, nevertheless the correlation coefficient between the other tests showed that comprehension and accuracy were related. These results support the story method of teaching suggested by Morris²² and, although memorized parroting by the child is to be deplored, methods suggested for teaching the application of intonation to reading involve memorization or quasi-memorization of the text.²³

The number of readers rated "non-fluent" in this investigation is in agreement with the extent of non-fluency found by Sheldon and Hatch in their grade two sample. If, as they suggest, non-fluency is common in beginning reading, then a short test is not likely to cause much change in the manner of reading. Fluency, the index of a reader with good comprehension of sentences and larger language structures was, with one exception, not apparent in the listening group. Nevertheless, the hypothesis was accepted since the test results suggest that listening to the

material read fluently with natural intonation helped the subjects in the listening group appreciate the structure of the passage so that not only comprehension was improved, but also reading accuracy and the identification of some new words from context clues.

The present study was short, and limited in its generality. Nevertheless, the results do suggest that hearing a model before making an oral response will, because the intonation patterns of speech have been heard, help the child grasp sentence meaning and identify words from context. In a longer study, the oral response which allows the child overtly to relate the paralinguistic signalling systems of speech to reading, may prove not only equal to, but more effective than, reading silently.

REFERENCES

- ¹ Edward L. Thorndike, "Reading as Reasoning: A Study of Mistakes in Paragraph Reading," *Journal of Educational Psychology*, Vol. VII (June, 1917), pp. 323-332.
- ² Charles C. Fries, *Linguistics and Reading* (New York: Holt, Rinehart and Winston, Inc., 1963).
- ³ Martin D. Braine, "On Learning the Grammatical Order of Words," *Psychological Review*, Vol. LXX (1963), pp. 323-348.
- ⁴ Donald J. Lloyd, "Intonation and Reading," *Education*, Vol. LXXXIV, (May, 1964), pp. 538-541.
- ⁵ David Gliessman, "Understanding Reading from the Viewpoint of Sentence Psychology," *The Reading Teacher*, Vol. XIII (1959-1960), pp. 22-28.
- ⁶ Walter F. Dearborn, P. W. Johnston and Leonard Carmichael, "Oral Stress and Meaning in Printed Material," *Science*, Vol. CX (1949), p. 404; Arthur I. Gates and E. B. Cason, "The Evaluation of Tests for the Diagnosis of Ability to Read by Phrases or 'Thought Units'," *The Elementary School Journal*, Vol. XLVI (1945-46), pp. 23-35.
- ⁷ B. F. Skinner, *Verbal Behaviour* (New York: Appleton-Century-Crofts, 1957).
- ⁸ O. Hobart Mowrer, *Learning Theory and the Symbolic Processes* (New York: John Wiley and Sons, 1960).
- ⁹ Authur W. Staats and Carolyn K. Staats, "A Comparison of the Development of Speech and Reading Behaviour with Implications for Research," *Child Development*, Vol. XXXIII (1962), pp. 831-846.
Vol. XXXIII (1962), pp. 831-846.
- ¹⁰ Cyrus D. Mead, "Silent Versus Oral Reading with One Hundred Sixth Grade Children," *Journal of Educational Psychology*, Vol. VI (1915), pp. 345-348.
- ¹¹ Charles E. Osgood, George J. Suci and Percy H. Tannanbaum. *The Measurement of Meaning* (Urbana: University of Illinois Press, 1957).
- ¹² Gertrude Hildreth, "The Role of Pronouncing and Sounding in Learning to Read," *Elementary School Journal*, Vol. LV (1955), pp. 263-269; Barbara A. Brinkley, et al., "Comprehension Following Silent and Oral Reading in Grades Two and Three," Unpublished Master's Thesis, Boston University, 1960; John D. McNeil and Evan R. Keislar, "Value of the Oral Response in Beginning Reading: An Experimental Study Using Programmed Instruction," *British Journal of Educational Psychology*, Vol. XXXIII (June, 1963), pp. 162-168.
- ¹³ Skinner, *loc. cit.*

- ¹⁴ Kenneth S. Goodman, "A Linguistic Study of Cues and Miscues in Reading," *Elementary English*, Vol. XXXII (October, 1965), pp. 639-643.
- ¹⁵ Muriel Potter Langman, "The Reading Process: A Descriptive Interdisciplinary Approach," *Genetic Psychology Monographs*, Vol. LXII (1960), pp. 3-40.
- ¹⁶ Mary H. Neville and Barry P. Frost, "Differential Achievement in Reading and Arithmetic," *The Alberta Journal of Educational Research*, Vol. X (December, 1964), p. 194.
- ¹⁷ Ruth G. Strickland, *The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children*, Bulletin of the School of Education, Indiana University, Vol. XXXVIII (July, 1962).
- ¹⁸ Albert J. Harris, *How to Increase Reading Ability* (New York: Longmans, Green and Co., 1961, 4th edition), p. 186.
- ¹⁹ William D. Sheldon and Shirley Hatch, "Strengths and Weaknesses in Reading of a Group of Second Grade Children," *Educational Administration and Supervision*, Vol. XXXVII (November, 1951), pp. 405-414.
- ²⁰ Robert A. McCracken, "The Oral Reading Performance of a Second Grade Class Using an Informal Reading Test," *Journal of Educational Research*, Vol. LV (November, 1961), pp. 113-117.
- ²¹ Brinkley, et al, *op. cit.*; McNeil and Keislar, *op. cit.*
- ²² Ronald Morris, *Success and Failure in Learning to Read* (London: Oldbourne, 1963).
- ²³ Donald J. Lloyd and Harry R. Warfel, *American English in its Cultural Setting* (New York: Alfred A. Knopf, 1956); Paul McKee, *The Teaching of Reading* (New York; Houghton Mifflin Co., 1948).

The marked trend toward subject specialization by teachers and the demand for higher academic qualifications implies that teaching behaviour and teaching technique are positively related to the qualifications of the teacher. The author of this study endeavoured to find significant differences between two groups of science teachers whose knowledge of subject matter determined by an achievement test differed significantly. A specialist in science teaching will provide a discussion of the problem and of this research in a subsequent issue of this journal.

PRISCILLA J. ECCLES

University of Alberta, Calgary

Knowledge of Science and Teaching Behaviour

Introduction

A number of studies of the relationship between teacher knowledge of subject matter and pupil achievement have been made. These were reviewed in a previous article¹ which also gave the inconclusive results of a similar study using sixth grade students and teachers from the public school system in Calgary.

This article reports observations of the teaching behaviour of a selected group of these teachers. Each had written the Iowa Tests of Educational Development, Test 2, General Background in the Natural Sciences, Form X-35² and for this series of observations the teachers with the top five scores and the lowest five scores were selected.

The intention was to compare these two groups of teachers in terms of what they did in the classroom as they went about teaching science. The overall hypothesis was that there would be observable differences in the behaviour of these two groups: one reasonably knowledgeable in the field of science, one less so.

The Sample

The following summarizes details given in the previously cited report. Thirty names were chosen at random from a list of sixth grade teachers in the Calgary Public School System. When a teacher was unwilling to take part another name was chosen. Thirty-eight teachers were asked and twenty-nine accepted. Of the nine who were not willing

to participate three had reasons which, in the opinion of the writer, were substantial and unrelated to any feeling of adequacy or inadequacy in the science field. The mean raw score on the Iowa Test of General Background in the Natural Sciences of these teachers was 45.9 (maximum possible 60) with a S.D. of 5.27. The range was from 35 to 58.

Extreme cases were selected for observational study, the five top scores and the five lowest. When these teachers were determined the range of scores for the lower group was 35-41 with a mean score of 38.6, but of this group two declined further participation. One other was not teaching science in 1962-63. Two accepted. Three teachers with the next highest score were contacted, one was not teaching in 1962-63 and one declined. Two more teachers were contacted, one of whom had excellent reasons for not participating, and one of whom accepted. One more contact completed the group. The range was finally 38-45 (standard scores 13-17) with a mean of 41.8. The difficulty experienced in obtaining agreement to participate must be noted and counted as a limitation of the study.

The range of scores for the top group was 51 to 58 (standard scores 19-21) with a mean of 53.8. One teacher was not teaching science in 1962-63 but it was possible to substitute another with a tied score so the range and mean remained unchanged.

These two groups of five teachers are referred to as the High and Low groups, although it will be realized that these are relative terms. There are no published norms for this test which would offer any guide for comparing these teachers to those in other provinces or states, but the mean score of the low group (41.8) is equivalent to a percentile rank of 84 on the twelfth grade norms for this test. It would be possible for a group of teachers to know much less about science than was the case for the five referred to as the "low" group. It should also be noted that the scores of the "high" group approached the upper limits of the test; some members may have a much more extensive knowledge than others.

The reader will also realize that in selecting for one trait such as knowledge of science, the possibility that one has selected for other traits as well is very great. It is a complex of traits resulting in, resulting from, or associated with knowledge of science that is being studied.

The groups appeared to be more comparable on other characteristics than one had a right to expect. For example, being appointed a principal indicates years of successful teaching and, normally, a University degree. In the High group three of the five were principals, but in the Low group two of the five were also principals. However, the High group did have some advantage in years of experience and in years of teacher education and might be expected to be more proficient classroom teachers. The teachers were not told the criterion for selection nor did they know their test scores.

Method of Observation

All the teachers used a timetable with two or three, thirty-to-forty-minute science periods a week. The times of the lessons were obtained and divided into ten-minute intervals. The design required ten visits to be made to each teacher and the ten-minute interval (i.e., the first ten minutes, the second ten minutes, the third ten minutes) was assigned randomly to each of these visits.

Ideally, the day of the visit should also have been a random selection, but some of the teachers taught science at an hour when the writer, too, had teaching commitments. It was necessary to visit these classrooms during the period classes were dismissed for student teaching and after the University term had ended in April. Each teacher was visited ten times, eighty-five observations were recorded on tape, of which three were not usable. Only seven recordings could be made in one classroom and, where necessary, the mean was used to provide an equal number of instances for analysis. For three teachers the final observation was not used in order that the scores reported for teachers would be based on the same number of observations. Table I gives the topics taught during the observations. These were all taken from the 6th grade content of the Alberta course of study and, with four exceptions, were all in the general area of the physical sciences. It would seem unlikely that, within this narrow range, the specific topic of the lessons could have had an appreciable effect on teaching behaviour. However, this assumption could be contested.

TABLE I
NUMBER OF OBSERVATIONS BY TOPIC

	Weather	Machines	Flight	Sound	Light	Other	Total
High Group	2	7	9	9	8	4	39
Low Group	8	11	9	3	7	2	40

No attempt was made during the observations to code or classify the behaviour observed. Using a watch with a second hand the writer described in twenty-second intervals the general classroom procedures and a small tape recorder was used to record verbal behaviour. No attempt was made to distinguish one student from another. On all occasions but one, the teachers used the entire class of thirty to thirty-five pupils as the instructional unit. When written work was assigned, and for very brief intervals at other times, teachers directed their attention to individual students. It was not possible to record the verbal interaction during these periods, which are described in the records as individual work. Seven per cent of the 790 minutes analyzed are in this category. In only one

sample was the class organized into small groups. The method of observation was not suited to this kind of instruction and this period was included in the above figure for individual instruction.

Given this kind of general teaching situation a single observer with a tape recorder can describe the type of activity, resources used, material put on blackboard, and the verbal behaviour of teachers and students. Only that part of the verbal behaviour of students which was intended to be part of the class activity was recorded but there was very little tendency for students to be overtly engaged in any activity other than that which was the appropriate general class activity at the time.

It was necessary to record some student comments in long-hand and occasionally comments were missed by both the observer and the tape recorder. Of all student comments intended to be heard by the class, nine per cent were not recorded. Most of these were probably not audible to many members of the class. The material on the tapes was divided into twenty-second intervals which were later broken down, by estimate, into five-second intervals, and typed on sheets with the descriptive material for that period. This material was classified under three headings:

1. Specific behaviours which may be associated directly with teacher knowledge such as failure to teach science, making factual errors, use of technical terminology.
2. Use of various teaching techniques.
3. Analysis of pupil comments in context.

Results

Many of the following results are given on an individual basis because, as will be seen, sometimes there were striking differences among individual teachers on behaviour variables which showed no difference for the two groups. Since this indicates that the behaviour observed was behaviour in which teachers do differ it emphasizes the importance of finding no relationship between this and subject matter knowledge.

In addition, there seemed to be *patterns* of behaviour, but with such small numbers it is not possible to determine if these are wholly a matter of individual style or if they might be related to general "characteristics" of teachers. Since these teachers were promised anonymity, the writer has deliberately obscured these patterns by being inconsistent in designating individual teachers by letter. This was done because the patterns that emerged from the data were, if nothing else, descriptive enough of a teacher's work so that he could recognize himself and very probably be recognized by others.

Specific Behaviours

Omitting Science Lessons. The most obvious effect of inadequate preparation for science teaching would be to omit teaching science. While

it is customary for teachers to schedule two or three periods a week for science, an elementary timetable is flexible. Within this flexibility a teacher who wishes to do so may avoid, consciously or unconsciously, a substantial portion of the lessons in a subject, especially if the subject is not a well-established one.

During the course of the study, there were occasions when science was not taught. On many of these the teacher volunteered the information that the lesson had been re-scheduled rather than omitted. No effort was made to inquire further or to elicit such statements, and no record was made of the volunteered statements. Attempts to judge the legitimacy of these reasons would have been out of order because of the possible adverse effect on the study itself, and because such inquiries, suggesting that the observer was questioning the teacher's veracity or judgement, would have been unethical. What was recorded was the simple fact that a science lesson was not being taught during the period scheduled and selected at random for observation. If there was a tendency for the Low group to do this for less valid reasons it should have occurred more frequently in their classrooms. But in considering this evidence, the reader should remember that any teacher who knowingly omitted science from the program would probably have been eliminated with the group who, for one reason or another, did not want to take part in the study.

On a number of other occasions the observer arrived, expecting to see a science lesson in progress, but found that it had not yet begun. When the visit was timed for the first segment of the lesson, no record of this was made since elementary timetables are not expected to operate on rigid schedule. But, if the sample was to have been from the middle or the last section of the lesson and the lesson had not started, or if the initial period was delayed more than ten minutes, a notation of "Late" was made. Again many of these lessons were merely late and nothing more and a full science lesson *may* have been taught, but there were times when the investigator thought that if she had not come there would not have been a lesson, at least that particular morning or afternoon. Again no judgments have been made as to what lessons these might have been, but if there was a greater tendency for the Low group to omit lessons the number of these late notations should have been larger for this group.

TABLE II
NUMBER OF UNTAUGHT AND LATE SCIENCE LESSONS

Group	Science not Observed	Late	Sub-total	Science Observed and on Time
Low	8	4	12	38
High	7	8	15	35

A chi-square test indicated this difference was not significant. These observations are not, strictly speaking, independent, and results are affected by personal idiosyncracies of the teachers. In this instance one member of the High group had a marked proneness to omit or begin science lessons late. Of the fifteen cases recorded, nine were in the classroom of this one teacher. But, even with this teacher removed from the sample, the difference was not significant.

Errors. Another gross effect of poor subject matter background would be errors—attempts to teach subject matter that is wrong. There were almost no simple mis-statements of facts recorded. As used here the term has a broader meaning including what might be more accurately called moments of confusion such as uncorrected slips of the tongue or use of an analogy that was, in the investigator’s opinion, unsound. It even includes failure to refine concepts to what would, again in the investigator’s opinion, be regarded as normal sixth grade understanding or to the degree implied by the objectives of the lesson. Examples of this are: failure to clarify at all a child’s confusion between X-rays and light; failure to make a distinction between myth and history in the history of flight. Such instances of “failure to” are specific omissions directly related to the content of the lesson in progress and do not refer to what the wirtter thought the teacher might have included but did not.

Even with such a broad definition, error was not frequent and always of short duration, that is the time during which the error affected the content of the lesson (statements, questions) was brief.

TABLE III
NUMBER OF ERRORS IN EIGHT 10-MINUTE OBSERVATIONS

Group	Lessons without Errors	Lessons with Errors
High	33	6
Low	32	8
The difference is obviously not significant.		

Use of Terminology. It seemed possible that more knowledgeable teachers might have a tendency to introduce concepts over and above those suggested by the course of study or to use technical terminology to a greater degree than less knowledgeable teachers.

Technical terminology was defined as words or short phrases, mastery of which was an objective of instruction, or words which, in the context of the lesson, referred to science concepts where some understanding was essential to make the lesson intelligible, (e.g., “light” in a lesson on mirrors was classified as a technical term but would not have been if it had been used in a casual comment such as “the light is bright today”).

The score for each lesson was the number of different technical words used. That is, a word such a “tornado” was counted once in a ten-minute observation regardless of frequency of use. It would have been recorded again only if it had been used during another ten-minute observation. This seemed to result in a reasonable estimate of the importance of technical words in the lesson, especially since the teachers would all cover the same subject matter during the year.

This data satisfied the conditions for Analysis of Variance and a group-within-treatments analysis was made, “treatments” referring to whether the lesson was taught by a High or Low group teacher, “groups” referring to the lessons taught by a single teacher.

TABLE IV
ANALYSIS OF VARIANCE, TECHNICAL WORDS USED PER LESSON SAMPLE

	d/f	S.S.	M.S.S.	P.
Treatments (High, Low)	1	4.99	4.99	n.s.
G _w T (Teachers)	8	151.94	18.99	n.s.
S (lesson sample)	70	1155.02	16.50	

Teaching Techniques

This section considers the following: time devoted to teacher talk, the time spent lecturing or “giving” notes (i.e., dictating or having children copy material from the blackboard). The use of books for reference, the use of teaching aids, time spent on administrative tasks, testing time, time spent giving individual help and, perhaps the most important for science instruction, time devoted to demonstrations and experiments.

With the addition of student talk, described in section three of this report, these categories included all activity that was observed. There is substantial overlap such as would occur when a teacher talks as he demonstrates or as children use reference materials.

There were no significant differences in time given to administrative routines or in working with individual students and only two observations out of seventy-nine including periods of testing.

Use of Teaching Aids. This category included the use of films, filmstrips, tapes, still pictures and reading from textbooks. It does not include pictures, exhibits, books, etc. available in the room unless used by the class during the period of observation. Children did use some of these materials in the periods of individual study and this was not recorded unless the class as a whole was instructed to do so. There is also a separate category for the use of concrete demonstration materials and for reference work.

In this restricted sense teaching aids were used too infrequently to apply a statistical test. Two teachers in each group made no use at all of

such aids during the periods observed. Children read from textbooks on two occasions in the classroom of one teacher (L) and twice in the classrooms of (H) group teachers. Filmstrips were being shown on two visits to the classroom of a High group teacher, one to the classroom of a Low group teacher. One teacher was showing a movie on one occasion (L) and one used a tape recording on one visit (H). With this meagre data there was not even the suggestion of a difference between the two groups.

Use of Books for Reference. This data refers to occasions when all children were directed to this kind of activity. Many of the books used were textbooks, but they were used to look up relevant information, not as a reading assignment. There was also occasional use of reference materials during individual study periods but these are not included.

TABLE V
NUMBER OF SECONDS BOOK USED FOR REFERENCE DURING
EIGHT 600-SECOND OBSERVATIONS

High Teacher	Score	Low Teacher	Score
a	0	a	0
b	100	b	120
c	635	c	0
d	1380	d	0
e	0	e	180

A Mann-Whitney U Test indicated that the null hypothesis of the High and Low group being from the same population should be accepted.

Total Teacher Talk. This refers to the time during which the teacher was talking to the class.

TABLE VI
MEAN NUMBER OF SECONDS OF TEACHER TALK
PER 600-SECOND OBSERVATION SAMPLE

High Teacher	Score	Low Teacher	Score
a	421	f	525*
b	405	g	356
c	381	h	324
d	285	i	292
e	266	j	229

*Total included 1700 S. of "note copying."

When the three ten-minute periods of "note giving" via the blackboard were classified as teacher talk, this data satisfied the requirements for Analysis of Variance. A groups-within-treatments analysis was made: "treatment" referring to the lessons being taught by a teacher of the High or Low group, "groups" referring to the lessons by an individual teacher, "subject" referring to the individual lesson sample.

TABLE VII
ANALYSIS OF VARIANCE: TEACHER TALK IN EIGHT 10-MINUTE SAMPLES

T	d/f	S.S.	M.S.S.	F-Ratio	P
	1	51.19	51.19 MS _A	.09	n.s.
G _w T	8	4647.32	580.91 MS _{G_wA}	9.24	P < .01
S	70	4400.28	62.86 MS _{wG}		

As was the case with several of the behaviours described in this paper there were significant differences among teachers, although this was not related to the teacher being in the High or Low group.

Monologue and Note Giving. Monologue was arbitrarily defined as the teacher talking for more than 60 consecutive seconds. The total amount of time here was quite small and most of these periods were extended explanation or introductory material. Teachers do not lecture to sixth graders. However, they do with some frequency use a related technique where information, sometimes new, sometimes review, is written on the board and copied by the children. The children usually had notebooks open in front of them—entries were often made in these as an activity interspersed with discussion, or as a result of reference work. This activity was not described as note copying. The term was restricted to periods when the children copied material from the board, while the teacher was usually engaged in writing the material, or in walking around the room giving individual help. Because this technique seemed closely related to “information giving” via teacher monologue, the two are combined to give Table VIII.

TABLE VIII
TOTAL NUMBER OF SECONDS OF MONOLOGUE — NOTE COPYING
IN EIGHT 600-SECOND OBSERVATIONS

High Teachers		Low Teachers	
a	820	a	1425
b	915	b	3285
c	900	c	355
d	980	d	1665
e	65	e	75

Randomization test for two independent samples indicates significance at .15 level.

Demonstration and Experiments. Although there were a few occasions when children were asked to draw directly on experience to state generalizations or rules, no truly experimental work was observed. In most instances demonstrations with concrete materials served as visual aids in support of verbal exposition by the teacher. The term “demon-

stration” was interpreted very broadly to include any instances where real materials or three-dimensional objects were an integral part of the lessons. Sometimes these materials were merely shown to the class, but on other occasions there were elaborate demonstrations of simple machines and the like. Sometimes students were asked to demonstrate materials and objects were frequently passed around the class. Times shown indicate periods during which the attention of the entire class was directed toward the demonstration or object.

TABLE IX
DEMONSTRATIONS: NUMBER OF SECONDS IN
EIGHT 600-SECOND OBSERVATIONS

High Teachers	Scores	Low Teacher	Scores
a	0	f	0
b	0	g	560
c	260	h	580
d	420	i	740
e	680	j	1560

Mann-Whitney U test indicates no significant difference between groups.

Frequently, failure to use concrete materials is inferred to be related to poor subject matter preparation. Not only do the data as a whole fail to support these inferences, but the teacher in this sample, whose work was most characterized by use of demonstration, was in the L group. It should also be noted that while the definition of demonstration used here was very broad, and had to be in order for some of the teachers to have any score at all, the work in this classroom was not only more extensive in time, but was also a more significant part of the lesson. The children were not just shown objects, they watched demonstrations and discussed what they had observed. Work of a similar quality was seen in the classroom of one, but only one, of the High group teachers.

Student Comments

Pupils’ comments were categorized into a rather detailed system which seemed suitable and adequate for the data. It is not reproduced here because the total in many of the categories was quite small and no trends relevant to high and low groupings of teachers were suggested. A statistical analysis was made of the total number of comments, the number of comments reporting past or present observations, and the number of comments in combined categories labeled “restrictive” and “less restrictive” as defined below.

Number of Comments. The expectation that Low group teachers might not have an equal quantity of pupil participation was based on the assumption that this is very much under the control of the teacher, and one who felt weak in the subject matter might have a tendency to use an approach

such as textbook reading, lecturing, or note giving which would not involve them in discussion. The total number of comments for individual teachers is given in Table X.

TABLE X
NUMBER OF PUPIL COMMENTS IN EIGHT 10-MINUTE OBSERVATION
SAMPLES

High Teacher	Score	Low Teacher	Score
a	129	f	96
b	169	g	52
c	131	h	125
d	158	i	109
e	110	j	194

As indicated in the previous section on teaching techniques, “discussion” was the dominant class activity for almost all of the teachers. Many of the student comments were, however, very brief. The data for number of comments per observation period satisfied requirements of normality of distribution and homogeneity of variance for a groups-within-treatments analysis of variance; “treatments” referring to whether the lesson was taught by a High or Low group teacher; “groups-within-treatments” referring to the group of lessons taught by individual teachers and “subject-within-groups” referring to the lesson samples.

TABLE XI
ANALYSIS OF VARIANCE FOR NUMBER OF STUDENT COMMENTS IN EACH
LESSON

T (H or L)	d/f	S.S.	M.S.S.	F-Ratio	P.
	1	210.52	214.52 (MS _A)	1.09	n.s.
G _w T	8	1569.48	196.18 (MS _{G_wA})	2.27	p < .05
S	70	6000.40	85.83 (MS _{wG})		

There was a significant difference among teachers as to the number of student comments made during lessons but these differences were not related to the teacher’s being in the High or Low group.

Comments Based on Pupil Observations. As reported above, there was no evidence that Low group teachers made less frequent use of demonstrations than did those in the High group. However, one might expect such teachers to be less able to incorporate such experience into the context of the lesson. In addition, if the teacher will ask the right questions, children can report many out-of-school observations which contribute to understanding the concept being discussed. High group

teachers might be expected to make more use of this kind of student contribution, i.e., reports of observations present or past, asked for or volunteered. A summary of these is given below:

TABLE XII
TOTAL NUMBER OF "REPORTS OF OBSERVATIONS"
IN EIGHT 10-MINUTE OBSERVATIONS

High Teacher	Score	Low Teacher	Score
a	1	a	16
b	28	b	2
c	5	c	20
d	13	d	11
e	31	e	41

A chi-square test of significance was based on the number of lessons in which zero, 1-4, or 5 or more such comments were made. There was no significant difference between the High and Low groups.

Number of Restricted and Unrestricted Comments. "Restricted" refers to the proportion of comments which were student replies to questions from the teacher asking for a "right" answer. Such comments require the recall of previously learned material or the application of facts and principles in explanation, prediction, or the naming of examples. These are considered restricted by the necessity of being right, in the sense that the teacher will accept the answer, but also in the sense that it is the right answer. Almost without exception the answers accepted were those facts, principles and definitions which would be considered correct (at this level) by the scientific community. They are considered correct because they have been verified by appropriate scientific processes or, in the case of definitions, by consensus. This aspect of "rightness" may be assumed but was seldom, if ever, made explicit. The answer was considered right because the teacher said that it was right. In either instance *once the question is asked* in a situation where both parties recognize the propriety of the teacher asking and the obligation of the pupil to answer, the student's comment is restricted by the necessity of its being correct. Both students and teachers clearly accept this necessity. Presumably if the comment is not right the student is sorry and the teacher disappointed. As a matter of record these comments were almost always acceptable. Most teachers evidently prefer to ask questions of students who indicate they know the answer or whom the teacher expects to know the answer. Only occasionally did a teacher ask a question where this did not seem to be the case, and almost without exception such wrong answers as were given were handled kindly. The teacher would, wherever possible, mention first that part of the answer which was correct—or reasonably so— then comment on and modify the incorrect

portion of the response. Hopelessly wrong answers were often passed without teacher comment, the teacher merely calling on another student who was indicating that he knew the right answer. In other words most of these comments represent student participation in what might be called a standard teaching process of eliciting and reinforcing correct responses. This is not the process of inquiry nor did these teachers intend it to be.

Interest here centres on whether teachers with a better science knowledge are more likely to promote a discovery rather than a receptive mode of learning. While the former would not be free of restricted student comments, there would surely be a lower percentage. Comments considered less restricted were of the following nature:

- (a) Content-oriented pupil inquiries; asking for information or explanation.
- (b) Statements reporting observations, both voluntary and in answer to a teacher's question, including reports of present and of past observations. Also included were student suggestions of observations that could be made but there was only one entry in this category.
- (c) Comments giving a prediction, explanation or novel example providing the context did not *require* a correct answer. Teacher questions about application elicit such answers but differ in clearly calling for the correct answer rather than a speculation or an hypothesis. These were answers to "What do you think?" if the teacher seemed to really mean "What do you think?" and not "What do you know?" All voluntary comments were assumed to be truly speculative, at least the emphasis on correctness was re subject matter and not the teacher, although many teachers reply to such comments by giving an authoritative answer.
- (d) Conclusions; forming a conclusion from a number of facts derived from discussion or experimentation, either in reply to a request to do so, or volunteered.
- (e) Comments reporting or relating to thought processes, i.e., given in reply to "How did you?" or "How should you?" when the context was a reasoning process rather than a standard procedure.

In all of the above the student was thought to have been freer to say what he thought, what he saw, or to ask what he wanted to know.

Good pupil questions, attention to thought processes, hypothesizing, making observations, and drawing conclusions are the raw material of an approach to science that involves the processes of inquiry. But these elements must be structured if process is to emerge. In only a few lessons, in the classrooms of the two teachers mentioned above as making good use of demonstrations, did some such structure or pattern emerge and this

not consistently or for long periods of time. For all the other teachers these comments were not central to the structure of the lesson. Table XIII gives two scores for each teacher; the percentage of “restricted” and “less-restricted” student comments.

TABLE XIII
PERCENTAGE* OF RESTRICTED AND LESS-RESTRICTED STUDENT
COMMENTS

High Teachers	Restricted	Less- Restricted	Low Teachers	Restricted	Less- Restricted
a	42	25	f	38	24
b	61	22	g	42	33
c	48	21	h	40	49
d	65	17	i	61	16
e	29	54	j	50	32

*Percentage is of total comments minus unintelligible comments. Totals are not 100 per cent because some comments, such as procedural inquiries and statements, are not considered in either category.

A Mann-Whitney U test indicated no significant difference between H and L groups.

A Kruskal-Wallis One-Way Analysis of Variance by Ranks indicated no significant difference among teachers for either restricted or less-restricted comments per lesson sample.

Discussion of Results

There was no significant difference between High and Low group teachers for any of the variables considered. These included number of lessons missed, errors committed, and use of technical vocabulary. Also studied were use of classroom time for administrative purposes, the amount of teacher talk, the amount of student talk, the use of audio-visual aids, the use of demonstrations and experiences and the use of reference books. No significant difference applies to the intellectual activity of the students as revealed in their comments.

Recognizing the possibility of error in accepting the null hypothesis, the results suggest that teachers' knowledge of subject matter may not be related to teaching behaviour and to that extent not related to teaching effectiveness, the latter statement being based on the assumption that there can be differences in effectiveness only if there are differences in relevant teaching behaviour.

Because it is frequently the case for students of educational problems to feel that there were real differences between groups not revealed by the crude measurements used, it is in order to say that in this instance the writer did not feel this to be so. I do not think that any person seeing these teachers at work could have placed them with any accuracy in the High or Low group. The teacher with the highest score would probably have been placed by any observer in the High group as would the teacher

with the lowest score in the Low group, but one did not sense any difference between the two groups.

Common sense, nevertheless, insists that knowledge of subject matter is essential for the teacher. There are several possible explanations for failure to support this conviction with evidence.

- (1) The most obvious is that the teachers were not really very different, that the range between the High and Low groups was not sufficient to be reflected in teaching behaviour.

It would be possible for an elementary teacher to know much less science than did the Low group and there is no upper limit. Further, it seemed to the writer that for many of the criteria the teachers whose behaviour differed most markedly from that of the group as a whole were the two who had the lowest and the highest test score. This is obscured for the reader by the way these scores are reported, for reasons previously given. It was by no means consistent throughout and may be merely coincidence.

- (2) Another explanation is that, contrary to the opinion expressed above, the appropriate behaviours were missed and the measurements were too crude to reflect differences in behaviour.
- (3) The third explanation is more speculative, suggested by the facts in this report rather than supported by them. It is that these teachers employ a mode and style of teaching within which variation in subject matter knowledge will have little effect.

This should not be interpreted as meaning that these teachers were not concerned with subject matter. They were wholeheartedly engaged in teaching the facts, principles and definitions outlined in the curriculum guide; and they were conscientious teachers with qualities of intelligence and general education which enabled them to prepare their work thoroughly. When they entered the classroom they had mastered the content they intended to teach. They may not have had a profound knowledge of the subject, but they did know the subject matter of the lessons.

What more than this could a learned person do? He could change the course of study, which contains a great deal of trivial material. These teachers do not consider this to be part of the teacher's role, nor are they encouraged to do so. He could enrich the given course but this may not be appropriate for the majority of his students. He may need to spend less time in preparation, but this is not observable in the classroom. There is the possibility that he will respond differently to student questions, either by answering questions that other teachers cannot answer or by incorporating these into the lessons with greater skill or beyond this, he will develop an approach to science teaching directed toward pupil understanding of the processes of inquiry.

With this possibility in mind it is interesting that the general teaching style observed in this study was not one in which student questions play a very important part. In 800 minutes of classroom observation only 81 content-oriented pupil questions were recorded. Forty-four of these were asked in the classes of High group teachers and thirty-seven in the classrooms of Low group teachers. Greater knowledge of the subject, as measured by the test, did not, in itself, result in a style of teaching which encourages pupil questions. It seems likely that none of these teachers regard this as a very important aspect of their work or, if they do, none have been able to develop techniques that will encourage a large number of good questions. All were polite and would listen to even the most rambling and incoherent questions with patience. They did not refuse to answer or infer that the pupil was wasting time. One suspected they were pleased with questions as indicators of pupil attention and interest, but it is clear that to actively develop skills and attitudes of inquiry just is not part of their conception of the teacher's role.

It is difficult to express these ideas without giving the impression that these were not very good teachers. The writer would like to emphasize that this is not the impression that is intended. These teachers were not fumbling around. Most of them, especially the more experienced, had a well-developed style of teaching within which they performed with grace and precision. It is suggested that for the kind of teaching observed in this study, which is compatible with the course of study which emphasizes the acquisition of knowledge, with the expectations of the community and pupils, and, very likely, with the self-concept of the teachers themselves, and which is done very well indeed by the majority of the teachers observed, a broad knowledge of the subject is not very relevant.

REFERENCES

- ¹ P. J. Eccles, "Subject matter competence of teachers and the quality of science instruction," *The Alberta Journal of Educational Research*, Vol. VIII, No. 4 (December, 1962), pp. 238-245.
- ² *Iowa Test of Educational Development*. Edited by E. F. Lindquist. (Chicago: Science Research Associates, Inc., 1952.)

NOTES ON CONTRIBUTORS

DONALD B. BLACK is the Registrar Designate at the University of Calgary. He is a graduate of The University of Alberta and holds an Ed.D. from Washington. He has been employed by the Canadian Defence Board, Educational Testing Service and most recently was Professor of Educational Psychology at The University of Alberta. While there, he served as secretary to the Alberta Advisory Committee on Educational Research, and subsequently as its second Director. Currently, his research interests are in selection procedures and teacher perceptions of the teaching art.

OLIVER H. BOWN is Professor of Educational Psychology, Associate Director of the Personality Research Center and Associate Director of the Research and Development Center for Teacher Education of The University of Texas. He is a recipient of the American Personnel and Guidance Association Research Award, a past president of the Texas Psychological Association, director of independent research projects and has published in the areas of personality, teacher education, counselling and prediction of college performance.

PRISCILLA J. ECCLES is an Associate Professor of Curriculum and Instruction at The University of Calgary. She received her doctorate from the University of Illinois, and taught science for two years in Canada. This is the fourth paper in the general area of subject matter competence and elementary science teaching.

FRANCES F. FULLER is Associate Professor of Educational Psychology and co-ordinator of the Assessment and Counseling Division of the Research and Development Center in Teacher Education of The University of Texas. Reports of her research in personality, psychotherapy and teacher education have appeared in the *Journal of Counseling Psychology*, *Personnel and Guidance Journal*, *Japanese Readings in Counseling Psychology*, *International Journal of Religious Education* and other publications. She is author of a forthcoming chapter in the 1967 *Association for Student Teaching Yearbook*.

WILLIAM D. KNILL has been a Research Associate with the Institute for Community Studies at the University of Oregon, and Assistant Professor of Education at the University of Saskatchewan, and is at present an Associate Professor of The University of Alberta. The study reported here is based upon the research strategies suggested earlier by Dr. Knill in the chapter he contributed to the publication, "The Politics of Education" in *The Local Community* (Danville: The Interstate Publisher, 1964).

ALVIN E. KOWALSKI was employed by the Calgary Separate School System for four years and recently has been appointed principal of St. Stephen's Elementary-Junior High School. He has done research for the Department of Chemistry, University of Saskatchewan and the Department of Educational Administration, The University of Alberta. While at the latter institute, he published *A Survey of Educational Theses* (Edmonton: The University of Alberta Press, 1965).

MARY H. NEVILLE is working toward her Ph.D. in Reading at The University of Calgary. She is a graduate of the University of New Zealand, and has taught in New Zealand, England and Northern Nigeria as well as in Canada. In 1964 she published an article in this journal with Barry P. Frost entitled "Differential Achievement in Reading and Arithmetic".

HERBERT G. RICHEK is chief of the Research Reports and Publications Section of the Assessment and Counseling Division of the Research and Development Center in Teacher Education of The University of Texas. In addition to numerous research reports and collaborative work, he has published in the *AERA Review of Research* (1964) and *Psychology in the Schools*.

MICHAEL SCRIVEN is Professor of History and Philosophy of Science at Indiana University, currently on leave at the University of California at Berkeley. He graduated in mathematics and physics from the University of Melbourne, took an M.A. in the combined schools of mathematics and philosophy there and went on to Oxford for the Ph.D. He has taught at the University of Minnesota and Swarthmore College. He is on the Executive Committee of the Social Science Education Consortium, Inc., and has been director of projects for them on the evaluation of curricula and values in the social studies curriculum for some years. He has just published a book on the non-technical problems of philosophy—*Primary Philosophy* (McGraw-Hill, 1966).

JAMES G. SNIDER is an Associate Professor of Educational Psychology at The University of Calgary. He previously taught at Foothills College, California, and has been a Visiting Professor at McGill University in Montreal. He has published articles in various professional journals, including the *Journal of Clinical Psychology*, *Personnel and Guidance Journal*, and *Psychology in the Schools*. He has recently completed a book of readings on the semantic differential in psychology and has partially completed a book to be called *The Psychology of Ownership*.

